HAL
open science

Binding e sostituzione sensoriale nell'informazione tattile e uditiva

Aurora Rizza

To cite this version:

Aurora Rizza. Binding e sostituzione sensoriale nell'informazione tattile e uditiva. Psychology. Sapienza University of Rome, 2019. Italian. NNT : . tel-02336488

HAL Id: tel-02336488

https://u-paris.hal.science/tel-02336488

Submitted on 31 Oct 2019

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

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

Facoltà di Medicina e Psicologia
Dipartimento di Psicologia

Dottorato di Psicologia e Scienza Cognitiva
XXX Ciclo

Binding e sostituzione sensoriale nell'informazione tattile e uditiva

Dottoranda
Aurora Rizza Belardinelli
Matricola 1163432

Supervisore
Prof.ssa Marta Olivetti

Co-tutor
Prof. J. Kevin O'Regan

Difesa della tesi: 14 settembre 2018

INDICE

Introduzione. **pag. 4**

1. **Background teorico**

1.1 Udito e tatto: premesse. **pag. 7**
1.1.1 Cosa hanno in comune e in cosa differiscono tatto e udito. **pag. 7**
1.1.2 Cosa sappiamo del tatto: 'cosa' e 'dove' nel tatto. **pag. 10**
1.1.2.1 Definizione e aspetti generali del *binding* unisensoriale aptico **pag. 10**
1.1.2.2 *Binding* '*what*' e '*where*' nel sistema aptico: studi comportamentali **pag. 11**
1.1.2.3 *Binding* '*what*' and '*where*' nel sistema aptico: network neurali e finestra temporale **pag. 18**
1.1.3 Cosa sappiamo dell'udito: 'cosa' e 'dove' nell'udito. **pag. 20**
1.1.3.1 Definizione e aspetti generali del *binding* unisensoriale uditivo **pag. 20**
1.1.3.2 *Binding* '*what*' e '*where*' nell'oggetto uditivo: studi comportamentali **pag. 25**
1.1.3.3 *Binding* '*what*' and '*where*' nell'oggetto uditivo: network neurali e finestra temporale **pag. 28**

1.2 Premesse sul *binding* e *binding* audio-tattile. **pag. 33**
1.2.1 ***Binding*:** definizione e aspetti generali **pag. 33**
1.2.1.1 Il problema o i problemi del *binding*: classificazioni **pag. 33**
1.2.1.2 Origine e soluzione del *binding* **pag. 37**
1.2.1.3 Aspetti temporali del *binding* **pag. 41**
1.2.2 *Binding* multisensoriale o crossmodale: aspetti generali **pag. 43**
1.2.2.1 Studi comportamentali e psicofisici **pag. 44**
1.2.2 Integrazione crossmodale a livello neuronale **pag. 48**
1.2.3 Dibattito sulla modalità dominante **pag. 50**
1.2.3 *Binding* audio-tattile: aspetti generali **pag. 52**
1.2.3.1 *Binding* audio-tattile: studi comportamentali e psicofisici **pag. 52**
1.2.3.1.1 Influenza del suono sulla percezione tattile **pag. 52**
1.2.3.1.2 Influenza del tatto sulla percezione uditiva **pag. 55**
1.2.3.2 *Binding* audio-tattile: network neurali. **pag. 57**
1.2.3.3 *Binding* audio-tattile: aspetti temporali. **pag. 62**
1.2.3.4 Sinestesia audio-tattile acquisita: il caso SR **pag. 65**

1.3 Sostituzione sensoriale. **pag. 71**
1.3.1 Cosa ci rende ottimisti rispetto alla sostituzione sensoriale? Evidenze sperimentali a favore. **pag. 73**
1.3.2 Critiche e limiti della teoria della sostituzione sensoriale **pag. 74**
1.3.3 Sensory motor theory **pag. 75**
1.3.3.4 Progetto "Feelspeech": ascoltare tramite la pelle. **pag. 76**
1.3.4.5 Perché usare il tatto per veicolare l'informazione linguistica? **pag. 77**
1.3.4.6 Comprendere il linguaggio tramite le mani: il metodo di Tadoma. **pag. 79**
1.3.4.7. Dispositivi tattili per la comprensione del linguaggio: background storico. **pag. 79**
1.3.4.7.1 Dispositivi tattili per la comprensione del linguaggio: caratteristiche tecniche. **pag. 81**
1.3.4.7.2 Dispositivi tattili per la comprensione del linguaggio:il superamento dei limiti. **pag. 87**
1.3.4.7.3 Dispositivi per la comprensione del linguaggio: focus sul training. **pag. 88**

1. 4 Costruzione del percetto: dalla percezione alla rappresentazione della realtà. **pag.98**
1.4.1 Formato della rappresentazione mentale: modale vs. amodale **pag. 103**
1.4.2 Rappresentazione mentale: focus sulla rappresentazione spaziale. **pag. 105**

2. Contributi sperimentali **pag. 109**
2.1. Integrazione audio-tattile e temporalità. **pag. 109**
2.1.1. Studio sulle finestre temporali: detezione di incongruenze temporali in flussi ritmici. **pag. 110**
2.1.2. Conclusione sull'integrazione audio-tattile e temporalità. **pag. 116**

2.2. Integrazione audio-tattile con stimoli linguistici. **pag. 118**
2.2.1 Esperimento psicofisico sulla percezione delle sillabe "BA-VA". **pag. 118**
2.2.2 Studio pilota alla ricerca dell'effetto "McGurk" audio-tattile **pag. 126**
2.2.3 Conclusione in merito all'integrazione audio-tattile con stimoli linguistici **pag. 130**
2.2.4 Studio pilota con due toni. **pag. 131**

2.3. Sostituzione sensoriale: 'ascoltare' con il tatto **pag. 136**
2.3.1 'Ascoltare con il tatto', senza informazione uditiva **pag. 137**
2.3.1.1 Studio pilota con il linguaggio tattile **pag. 137**
2.3.1.2 Studio pilota con l'utilizzo della modalità tattile e lettura delle labbra **pag. 141**
2.3.2 'Ascoltare con il tatto', con parziale (o residua) informazione uditiva **pag. 145**
2.3.2.1 Studio pilota con l'audiolibro **pag. 145**
2.3.2.2 Esperimento con l'audiolibro **pag. 146**
2.3.3 Conclusione sulla sostituzione sensoriale 'ascoltare con il tatto' **pag. 149**

2.4. Rappresentazione della realtà: sopramodale o modalità specifica **pag. 151**
2.4.1 Studio con mappe. **pag. 151**

3. Discussione **pag. 156**

4. Conclusione **pag. 164**

5. Bibliografia **pag. 169**

6. Appendice **pag. 189**

Introduzione

In una serata estiva ventosa, durante l'esecuzione di uno spettacolo teatrale, la regia mette in sottofondo il rumore del vento. Improvvisamente la sensazione del vento che soffiava da tutta la serata, divenne molto forte, quasi fastidioso. Iniziai a sentire freddo. Mi soffermai ad analizzare la situazione e a domandarmi: è possibile che il vento sia aumentato in coincidenza con la proiezione audio, oppure che l'ascolto del rumore del vento abbia influenzato, e quindi, rafforzato la sensazione del vento sulla pelle a tal punto da sentire freddo?
Sempre nella stessa estate, seduta in un locale notai, sorseggiando una bibita, che la cannuccia poggiata al fondo del bicchiere vibrava in base alle frequenze della musica che veniva trasmessa nel locale. Ad un tratto, lasciando la presa della cannuccia, mi accorsi che il volume della musica del locale sembrava come 'diminuito', o meglio, con un volume inferiore rispetto a quando tenevo la cannuccia vibrante in mano.
Questi due fenomeni, il primo in cui una informazione uditiva ha influenzato la percezione tattile (in questo caso termica), e il secondo, in cui una informazione tattile ha influenzato la percezione uditiva del volume, hanno esperienzialmente confermato una parte delle ipotesi del presente lavoro.

Molti di voi sapranno per esperienza che una vibrazione o la percussione di un qualsiasi materiale comporta la produzione di un suono...ma forse non molti di voi avranno notato che la suoneria del cellulare sembra ad un volume più alto se concomitante alla vibrazione, rispetto alla medesima impostazione della suoneria senza la vibrazione. Se la cosa vi sembrerà insolita, basti pensare alla situazione descritta da Schürmann e coll. (2004), in cui suoni presentati simultaneamente ad una stimolazione vibratoria vengono percepiti ad un volume più alto, rispetto a quando vengono presentati isolatamente.

Per tanti anni i sensi sono stati studiati in modo ‘isolato’, come se l’uomo fosse una macchina di elaborazione a compartimenti stagni. Di certo una tinta di colore viene esperita attraverso la vista, un pizzicotto viene avvertito attraverso il tatto e un suono potrà essere differenziato solo grazie all’udito, ma nel processo di rappresentazione del mondo circostante i sensi si influenzano l’uno con l’altro. Il mondo che ci circonda, quindi, è frutto di una mutua influenza di una modalità sull’altra, nonché una rappresentazione mentale dell’unione dei sensi.

Nella parte introduttiva di questo lavoro sono stati descritti i sensi del tatto e dell’udito, per quello che di loro conosciamo e presentati i punti in comune di queste due modalità. Udito e tatto non solo condividono delle proprietà psicofisiche, ma anche elaborano le informazioni in ingresso in modo similare, ovvero analizzando la struttura temporale (sequenziale) degli stimoli (il che, per esempio, le differenzia in modo sostanziale dalla modalità visiva).

Proprio grazie a queste somiglianze, tatto e udito sono degli ottimi candidati per studiare fenomeni percettivi quali l'integrazione crossmodale. Paradossalmente, in letteratura sono presenti molti studi sull'integrazione audio-visiva e visuo-tattile, ma poco è stato detto in merito all'integrazione della modalità tattile e uditiva. Ho quindi riportato le principali ipotesi teoriche formulate in merito al problema generale dell'integrazione crossmodale, con focus sull'origine e soluzione dei problemi legati al '*binding problem'* o per meglio dirsi, '*binding problems'* e analizzato la letteratura sull'integrazione cross-modale audio-tattile. In particolare si è fatto riferimento agli studi comportamentali e di neuroimmagine, cercando di capire i meccanismi con cui tatto e udito interagiscono tra loro e come si influenzano vicendevolmente.

L'unico caso di sinestesia audio-tattile presente in letteratura, conseguente a lesione cerebrale talamica, il che fa supporre un legame tra queste due modalità a livello cerebrale (attiguità neuroanatomica) e supporta le teorie sulla plasticità neuronale.

Dal momento che stimoli tattili presentano diversi punti in comune con gli stimoli uditivi, ci si è posti la domanda se sia possibile sentire attraverso il tatto. La teoria della sostituzione sensoriale afferma che è possibile veicolare informazioni di una modalità sensoriale attraverso un'altra modalità. Sebbene, infatti, ogni organo di senso codifichi gli input esterni con la specificità che lo riguarda, il sistema di conversione degli stimoli (in input elettrici), rende parte di questo contenuto informativo trasferibile e codificabile anche da modalità sensoriali differenti. In altre parole, vi sono delle proprietà degli input sensoriali che sono comuni ai diversi sensi, da cui consegue l'esistenza di una flessibilità funzionale alla base delle diverse modalità sensoriali.

La seconda parte del lavoro, riporta alcuni contributi empirici che sono stati svolti durante il periodo di dottorato, per rispondere alle domande aperte che sono emerse dal lavoro di ricerca bibliografica. In particolare, dal momento che la letteratura non riportava studi sulle finestre temporali audio-tattili per flussi di stimoli ritmici continui, sono stati svolti due studi unisensoriali (modalità uditiva e tattile) e uno studio multisensoriale audio-tattile per indagare la finestra temporale di flussi audio-tattili, in cui questa integrazione si verifica. Si è indagato inoltre, come riscontrare l'integrazione audio-tattile in modo automatico, come percezione pura, utilizzando stimoli uditivi linguistici semplici. In particolare si è indagata l'influenza della stimolazione tattile sulla percezione di sillabe con un esperimento psicofisico (esperimento sulla discriminazione delle sillabe 'ba' e 'va') e uno studio pilota ispirato dal 'McGurk effect' audio-visivo.

Riguardo alla teoria della sostituzione sensoriale, si è indagata la presenza di una facilitazione della modalità tattile sulla percezione uditiva di stimoli linguistici con la speranza

di poter sviluppare un sistema di sostituzione sensoriale in grado di aiutare le popolazioni con deficit uditivi a integrare le informazioni uditive mancanti con l'informazione tattile. Si è svolto uno studio in presenza di stimolazione uditiva degradata (esperimento con un audiolibro che veicola informazione degradata), ed un pilota in completa assenza di stimolazione uditiva (apprendimento e test di comprensione linguistica).

I risultati ottenuti hanno corroborato i dati precedenti sulle finestre temporali e l'ipotesi dell'esistenza di una rappresentazione spaziale sovramodale. Per quanto riguarda il fenomeno di integrazione audio-tattile e la teoria della sostituzione sensoriale, i risultati non hanno corroborato l'esistenza di un fenomeno di integrazione audio-tattile puro per stimoli linguistici, né una sostituzione sensoriale uditiva con la stimolazione tattile. Gli studi sono stati analizzati uno per uno e sono state poste nuove domande di tipo metodologico per lo studio dei processi di integrazione e di sostituzione sensoriale.

Scopo del presente lavoro di ricerca è non solo la conoscenza dei processi di elaborazione audio-tattili della popolazione umana, ma soprattutto il loro impatto applicativo molto importante in ambito riabilitativo in soggetti con deficit sensoriali. Inoltre, così come per lo studio dell'interazione cross-modale in generale, si reputa che lo studio dell'integrazione audio-tattile possa contribuire sia a chiarire alcuni aspetti dell'elaborazione cognitiva sia il controllo dell'azione (Meredith and Stein 1983, 1985; Stein and Meredith 1993; Macaluso and Driver 2005; Kayser and Logothetis 2007).

1. BACKGROUND TEORICO

1.1 Udito e tatto: premesse.

Cosa hanno in comune e in cosa differiscono tatto e udito.

In questo lavoro si è deciso di prendere in esame tatto e udito, nella loro interazione e mutua influenza, per diversi motivi. Come accennato, nella vita quotidiana non sono pochi gli esempi in cui tatto e udito interagiscono tra loro e diversi studi hanno mostrato una sistematica interazione percettiva tra queste due modalità (si veda paragrafo sul binding audio-tattile). Una delle ipotesi per cui questa interazione sia così comune potrebbe dipendere dal semplice presupposto che queste due modalità presentano notevoli similarità (Ro et al. 2013).

Fisiologicamente, la similarità tra l'elaborazione uditiva e tattile inizia dalla struttura e funzione degli organi recettivi. Entrambi infatti, dipendono da uno spostamento meccanico (le stereociglia delle cellule pilifere cocleari per l'udito; e i meccanorecettori della membrana cellulare per la vibrazione della pelle) per codificare eventi fisici in segnali elettrici o neurali. Per entrambi, inoltre, la codifica viene suddivisa in 'fasce' di frequenza, per cui ad ogni range di frequenze corrispondono differenti popolazioni di stereociglia o meccanorecettori che codificano per diverse frequenze.

Nel suo lavoro pionieristico, von Békésy (1955, 1957, 1959) ha studiato in modo sistematico le similarità (puramente fisiche, nella trasmissione delle onde/vibrazioni) tra i sensi del tatto e dell'udito, a tal punto che l'autore considera il tatto come un modello affidabile per lo studio delle caratteristiche funzionali dell'udito (Gescheider, 1970). In questo lavoro l'autore riprende la teoria della trasmissione delle onde e il modello meccanico di funzionamento della membrana cocleare. Il punto di partenza degli studi svolti è quindi fisico, dopo di che vengono tratte le supposizioni sul funzionamento.

L'analogia dei meccanismi di codifica a livello delle superfici dei rispettivi recettori (pattern di deformazione simile tra le membrana basilare dell'orecchio interno e i meccanorecettori della pelle), rispondono allo stesso tipo di energia fisica (pressione meccanica ad una specifica vibrazione di onde) (si veda von Békésy, 1955).

Le analogie nei meccanismi fisiologici (si veda Corey, 2003; Gillespie & Müller, 2009) sottolineate dalla percezione vibrotattile e uditiva sembrano condividere origini comuni (per delle rassegne in merito si veda Soto-Faraco & Deco, 2009; and von Békésy, 1959). Lo stesso von Békésy ipotizza che queste similarità siano dovute in parte alle origini evolutive della membrana basilare come conseguente ad una evoluzione di un'area della pelle che nel tempo è diventata estremamente sensibile alle vibrazioni, come accade in alcune specie di pesci nella parte chiamata linea laterale, si veda Montgomery e Macdonald, 1987). Il legame tra udito e tatto è stato quindi più volte dimostrato (si veda Marks, 1983; von Békésy, 1959).

A livello neuroanatomico, l'area primaria uditiva (A1) è collocata nel piano temporale superiore che è adiacente alla corteccia somatosensoriale secondaria nell'opercolo parietale (si

veda il caso di sinestesia audio-tattile riportato da Ro et al., 2007). Questo lascia supporre che, a livello neuroanatomico la loro integrazione sia plausibile. Inoltre, si è visto che regioni della corteccia uditiva umana sono reattive al tatto (Foxe et al. 2002; Kayser et al. 2005; Schurmann et al. 2006) e che la corteccia somatosensoriale secondaria risponde ad una stimolazione uditiva (Ro et al., 2013) supportando l'esistenza di una loro connessione (Beauchamp and Ro, 2008). In verità questi ultimi studi suggeriscono questa ipotesi, ma non indagano direttamente le connessioni anatomiche (nello studio di Ro et al. 2013 viene fatta una tractografia deterministica e probabilistica del peso diffusione in MRI che esamina la densità delle fibre di connessione) dal momento che risulta metodologicamente complessa.

Tra le differenze tra tatto e udito, von Bekesy riporta la diversa temporalità delle due modalità: la trasduzione e sensibilità di soglia assoluta dell'orecchio rispetto alla pelle è più rapida (dovuta alla maggiore densità di recettori del sistema uditivo), il che produce una latenza inferiore dalla stimolazione alla sensazione completa dell'udito rispetto al tatto. Inoltre, le due modalità presentano un diverso range di frequenze di vibrazioni a cui sono sensibili. La pelle, infatti, presenta la capacità di registrare oscillazioni molto piccole perchè è sensibile alla pressione meccanica di natura non vibratoria], le oscillazioni hanno bisogno di una frequenza al di sopra dei 3Hz. Di contro, il sistema uditivo è sensibile alle frequenze all'interno del range da circa 20 Hz ai 20 kHz, in popolazione di giovani adulti (Coren et al, 1999). Malgrado queste differenze, bisogna notare che anche se piccola, esiste una finestra di sovrapposizione della sensibilità delle vibrazioni, il che fornisce una base per l'interazione audio-tattile di natura integrativa già ad uno stadio precoce di elaborazione dell'informazione.

Inoltre, il tatto presenta degli aspetti meccanici (non considerando gli aspetti termici e sensazioni di dolore) che sono tradotti nella periferia da quattro tipi di recettori meccanici, ognuno con una specifica morfologia, range di sensibilità, dimensione del campo recettivo, e la risposta nel tempo (si veda Jones et al., 2006 per una rassegna). Di conseguenza l'attivazione selettiva di recettori diversi comporta una sensazione diversa (per esempio, la pressione costante dall'attivazione di Merkel e Ruffini, o la vibrazione o formicolio, pizzicore in seguito all'attivazione dei corpuscoli di Meissner e Pacini). Tutti e quattro i tipi di recettori sono comunemente coinvolti, con gradi differenti, nell'esperienza aptica di contatto con gli oggetti e le superfici. D'altro canto, invece, la trasduzione del suono non implica una varietà così ricca di recettori, con le cellule pilifere attaccate alla membrana basilare come base della trasduzione dell'informazione acustica (Moore, 1997). In sintesi, mentre il sistema tattile presenta diverso tipo di recettori, per il sistema uditivo la trasduzione è molto più semplice.

Un’altra distinzione tra tatto e udito può essere fatta considerando l’organizzazione di ogni senso nel sistema nervoso centrale. A livello di organizzazione funzionale delle cortecce sensoriali, mentre l'organizzazione corticale somatotopica presenta un arrangiamento spaziale somatotopico (Penfield, 1950), la corteccia uditiva è organizzata in regioni distinte che contengono neuroni che sono massimamente sensibili a determinate frequenze (ovvero un'organizzazione di tipo tonotopico, per così dire (Phillips, 1993). Ragion per cui per l’udito, gli stimoli vibratori riflettono la localizzazione dei neuroni reattivi della superficie corticale della specifica ‘suddivisione’ della corteccia uditiva. Per completezza, un’eccezione a questa organizzazione ‘tonotopica’ viene riscontrata nelle frequenze tra 5Hz e 50Hz. Inoltre, sembra che proprio entro questo range di frequenze, il contenuto della frequenza vibratoria sulla pelle così come nell’orecchio, sembra riflettersi nel ritmo di scarica dei neuroni nelle rispettive cortecce sensoriali (per esempio Romo et a., 2001 e 2003 per la corteccia somatosensoriale e per la corteccia uditiva Bendor et al. 2007 e Lu et al. 2001, entrambi studi con primati non umani).

In conclusione, la stretta relazione tra il tipo di codifica degli eventi vibrotattili della pelle e dell'orecchio (contenuto dell'informazione in termini di pressione meccanica tradotta in pattern oscillatori) e il range di sensibilità delle vibrazioni che si sovrappongono per entrambe le modalità, supportano una naturale integrazione delle due modalità (dagli stadi periferici fino agli stadi delle aree corticali associative del sistema nervoso centrale). Ne consegue che queste due modalità sono delle candidate a processi di integrazione audio-tattile (si veda capitolo sul *binding* audio-tattile) che fanno supporre che esistano dei meccanismi di integrazione multisensoriale tra udito e tatto per avere informazioni sull’ambiente circostante (si pensi al suono emesso dallo strofinio della mano su una tessitura). Esempi di interazione audio-tattile sono: suoni a determinate frequenze possono alterare la percezione di vibrazioni alle stesse o simili frequenze (Ro et al. 2009; Yau et al. 2009; Wilson et al. 2010), e viceversa (Yau et al. 2010), stimoli tattili possono influenzare la percezione del linguaggio (Gick e Derrick 2009; Ito et al. 2009) e viceversa (Ito e Ostry 2012), frequenze uditive alte e intense possono alterare la percezione di tessiture (Jousmaki e Hari 1998; Guest et al. 2002).

Esistono inoltre, diverse ipotesi, in merito al fatto che l’integrazione audio-tattile in qualche modo sia facilitata, dal momento che le vibrazioni tattili e sonore sono tra i primi stimoli elaborati dal feto (più precisamente il feto risponde a questi due tipi di informazioni a partire dal terzo trimestre Kisilevsky, 1995). Infatti, la sincronia temporale audio-tattile è la prima forma di elaborazione crossmodale (Lewkowicz, 2000). È quindi plausibile che la presentazione degli stimoli in prossimità temporale possa supportare lo sviluppo successivo

dell'elaborazione degli input sensoriali crossmodali che si verificano simultaneamente (si veda Lecaunet & Schaal, 1996).

1.1.2 Cosa sappiamo del tatto: 'cosa' e 'dove' nel tatto

Definizione e aspetti generali del *binding* unisensoriale aptico.

Il termine aptico deriva dal greco *haptós*, palpabile, *haptikós*, adatto al tatto.

Citato da da Revesz (1950) come la capacità di afferrare, il termine 'aptico' ingloba storicamente: a) il termine *cinestesi*, coniato da Bastian (1880) come senso del movimento nel riconoscere la complessità del contributo di pelle, muscoli e tendini nel generare l'impressione di sentire il peso, la resistenza, lo sforzo, la posizione e i movimenti dei lembi (Boring, 1942), per certi versi legato all'introspezione (Bastian, 1880); b) il termine di 'propriocezione' (Sherrington, 1906) intesa come percezione di sé da un punto di vista più fisiologico che introspettivo.

Lee (1978, 1980) aggiunge il termine 'expropocezione' per enfatizzare il ruolo dell'ambiente rispetto all'organismo.

Successivamente, Shaw (2001) riprende il concetto di expropocezione e conia il termine proexterocezione, relativo alla percezione del sè relativo all'ambiente. Egli infatti vuole indicare che implicitamente il termine aptico include anche lo 'scopo' nel fare qualcosa e potrebbe essere incluso in una teoria dell'intenzionalità (Shaw, 2001).

Carello e Turvey (1995, 2004, 2011) sintetizzano le terminologie: exterocezione (lunghezza di un oggetto), expropocezione (l'inclinazione di un oggetto nella mano), propocezione (disposizione della mano), proexterocezione (localizzazione della mano rispetto ad un oggetto) (vedere fig.1, Turvey e Fonseca 2014).

Turvey e Fonseca (2014) ipotizzano che alla base del senso aptico vi sia la tensegrità, intesa come struttura geometrica più adatta a spiegare il senso aptico. Gli autori affermano che la chiave delle funzioni del sistema percettivo aptico è la deformazione temporanea del tessuto. Gli autori si focalizzano sull'analisi del tessuto connettivo e la sua congiunzione con I muscoli e lo scheletro, presupponendo che ogni cellula dell'organismo è un sistema aptico. Il corpo umano risponde ad una specifica deformazione del sistema per cui, per esempio la pressione uniforme non è deformante e quindi una stimolazione non meccanica. Lo 'strumento di trasmissione' o meglio, il *medium* biofisico del sistema aptico deve essere omogeneo e isoforme così come lo sono l'aria e l'acqua (mezzi di trasmissione per eccellenza, necessari per gli altri sistemi sensoriali: la compressibilità permette di trasmettere eventi meccanici, la

trasparenza la trasmissione della luce, la diffusività permette le interazioni chimiche). Ovviamente gli autori parlano di invarianza all'interno di una eterogeneità (Elsasser, 1998) tipica di un sistema che per definizione è 'individuale'.

In conclusione, come riportato da Lederman e Klatzky (2009), l'aptico è un sistema percettivo mediato da due sottosistemi afferenti, cutaneo e cinestesico, che tipicamente coinvolgono un'esplorazione manuale attiva. Questo sistema consente soprattutto l'elaborazione di caratteristiche materiali di superfici e oggetti. Il sistema aptico usa l'informazione sensoriale che deriva da meccanorecettori e termo recettori della pelle (input cutanei) insieme ai meccanorecettori dei muscoli, tendini e giunture (input cinestetici). Gli input cutanei e cinestetici sono combinati e pesano in diverso modo per le varie funzioni aptiche.

Il sistema cutaneo ha una limitata capacità di risoluzione spazio-temporale nel rilevare dettagli degli oggetti (Lederman e Klatzky, 2009). In alcune circostanze, questi fattori condizionano la percezione aptica. Una vasta porzione della letteratura ha studiato la soglia di discriminazione di due punti, ovvero quale distanza minima tra due punti consente di percepirli come distinti o uniti (il soggetto percepisce un singolo punto). Si tenga in considerazione che questi studi prendono il via dal ricercare il corrispettivo 'recettore' ritrovato in ambito visivo, per cui si suppone che se si percepiscono due punti distinti allora ciascun punto attiva un campo recettivo diverso (Weinstein, 1968). Altri studi psicofisici in merito al limite del sistema cutaneo sono svolti chiedendo al soggetto se uno schema graticolare lineare viene appoggiato verticalmente o orizzontalmente (Johnson e Phillips, 1981; Craig, 1999). I risultati di questa parte della letteratura (studi comportamentali e poi delle basi neurali, vedi "*Homunculus*") hanno dimostrato che la sensibilità tattile è diversa per ogni parte del corpo.

James e coll. (2007) distinguono il tatto, come senso passivo dall'aptico, inteso come senso attivo che implica un atto esplorativo.

Binding '*what*' e '*where*' nel sistema aptico: studi comportamentali.

Uno dei primi autori che ha parlato del sistema aptico è stato J. J. Gibson nel 1962. Dalla definizione originaria di "sensibilità dell'individuo verso il mondo adiacente al suo corpo", la definizione di sistema aptico è stata legata all'esplorazione attiva degli oggetti. Gibson enfatizza la polarità dell'esperienza aptica dell'essere passivamente toccati, il che conduce l'attenzione dell'osservatore sulle sensazioni corporee soggettive e il contatto risultante da un'esplorazione attiva, che tende a guidare l'attenzione dell'osservatore verso le proprietà dell'ambiente esterno. Se vogliamo, nella definizione stessa del senso aptico è possibile vedere

unite in modo imprescindibile, le caratteristiche del '*what*' e '*where*' accennati precedentemente per il sistema uditivo.
Lederman e Klatzky (2009) si riallacciano esplicitamente ai due sistemi del 'cosa' e del 'dove' riportati in letteratura: il 'cosa' ha a che fare con le funzioni percettive (e mnestiche), e 'dove' viene riferito alla guida percettiva dell'azione.

Il sistema del **'cosa'** nel tatto elabora superfici, oggetti e molte altre loro proprietà, L'efficacia di questa elaborazione è dimostrata dalla scoperta che oggetti familiari sono riconosciuti più rapidamente e con più accuratezza rispetto al tatto da solo (Klatzky, Lederman, & Metzger, 1985). Questa abilità ha origine dai segnali sensoriali primitivi dei recettori periferici. Un ampio spettro di proprietà risulta da un'ulteriore elaborazione neurale dei segnali dei recettori. Secondo Lederman e Klatzky (2009) è utile distinguere le proprietà degli oggetti accessibili apticamente in due ampie classi: materiale e geometrica. Le proprietà materiali sono definite come quelle indipendenti dallo specifico campione di oggetti che viene considerato, mentre le proprietà geometriche descrivono la struttura di quel campione di oggetti. Le principali proprietà materiali sono: la tessitura di superficie, ruvidità, durezza, qualità termica, elasticità e temperatura. Le principali proprietà geometriche sono la forma e la dimensione. Il peso è una sorta di proprietà ibrida che riflette sia una proprietà materiale (es. densità) che della struttura (es. volume). Gli studi in tal senso sono molti e variano per paradigma e oggetto di studio. Per esempio, la tessitura di superficie può essere caratterizzata dalla ruvidità, viscosità, scivolosità, frizione, ecc. o la dimensione può essere calcolata in base all'area totale, volume, perimetro, ecc.
Lederman e Klatzky (1987) riportano sei procedure di esplorazione manuale, o "procedure esploratorie", associate alle proprietà degli oggetti in seguito alla loro esplorazione attiva. Gli autori hanno chiesto ai soggetti di esplorare le proprietà di alcuni oggetti e da questo studio hanno tratto le seguenti modalità esplorative: 1) movimento laterale per l'analisi della tessitura; 2) "proprietà non dimostrata" ovvero il sostenere un peso con il palmo aperto; 3) pressione per verificare la durezza; 4) presa chiusa per esaminare la forma globale e il volume; 5) contatto statico per la temperatura; 6) seguire il contorno per la forma globale e precisa. L'ipotesi degli autori è che ognuna di queste procedure è caratterizzata non solo dall'azione motoria ma anche per l'attivazione neurale (recettori neurali associati) che a sua volta è collegata (e facilita) determinati meccanismi computazionali. Ogni procedura esplorativa ha determinati costi e benefici (Klatzky e Lederman, 1993; Lederman e Klatzky, R. L.,1987) ed in sintesi, gli autori

ritengono che le strategie migliori in assoluto (minore costo, maggiore beneficio) sono il sostegno del peso, il contatto statico e la presa chiusa.
Come accennato, ridurre il sistema aptico al solo sistema somatosensoriale, non aiuta a comprendere la complessità del fenomeno, anche da un punto di vista neurale. Ciò è anche ovvio se si considera che una singola prensione di un oggetto, non è sufficiente per identificarlo (Bodegård et al. 2001; Roland and Mortensen 1987; Seitz et al. 1991), mentre per identificarlo occorre integrare appunto, l'input della pelle con l'informazione data da muscoli e giunture (componente cinestetica). Klatzky and Lederman (1995) indicano che una breve esposizione aptica (chiamata "*haptic glance*") è sufficiente per identificare un oggetto già conosciuto. Ai soggetti viene consentito di toccare gli oggetti (senza muovere le dita) per 200 ms, il che è in linea con il tempo necessario per il riconoscimento di un oggetto utilizzato anche da Gurtubay-Antolin e coll. (2004) e con lo studio di Klatzky e coll. del 1985, in cui veniva richiesto di nominare degli oggetti lasciando libera l'esplorazione tattile.
Tra le proprietà più studiate degli oggetti vi sono la forma e la tessitura (James e coll., 2007). Una delle ragioni per cui la forma (proprietà geometrica insieme alla dimensione) è stata molto studiata è legato al fatto che la forma è sufficiente di per sé per riconoscere un oggetto (Biederman, 1987) se si ha come scopo quello di risconoscerlo (Klatzky, Lederman e Reed, 1987) e di categorizzarlo come appartenente ad una classe di oggetti (Lederman e Klatzky, 1990).
Come per la visione, l'acquisizione dell'informazione aptica, parte dalla attivazione dei recettori di superficie più o meno acuti (la parte con recettori più acuti sono le dita della mano) e come vedremo, l'esplorazione e come detto, i movimenti delle dita sono più o meno stereotipati e 'vengono scelti' in base alle proprietà che si vogliono analizzare, materiali o geometriche. Ogni procedura esplorativa richiede un tempo ben preciso e viene eseguita in modo sequenziale.

Il sistema del '**dove**' per il tatto fornisce la descrizione dello schema dei punti, superfici e oggetti nel mondo. Il tatto differisce dalla visione, perché la quale la localizzazione può essere riferita all'organo sensoriale stesso, la pelle, così come all'ambiente. Dunque, gli autori distinguono due tipi di sistemi di riferimento per la localizzazione spaziale aptica: quello che determina dove lo stimolo viene applicato sul corpo, e quello che determina dove lo stimolo è stato toccato nello spazio esterno al corpo, senza dimenticare i punti di riferimento egocentrico e allocentrico (in base alla distanza e direzione; relativi all'attore della sensazione o legati a punti di riferimento esterni, rispettivamente).

Localizzazione spaziale sul corpo in condizioni passive di contatto (risposta alla domanda "Dove sono stato toccato?"). L'accuratezza della localizzazione corporea dipende da diversi fattori tra cui: a) la sensibilità della pelle che varia in base alla zona corporea stimolata, età ed esperienza visiva (strettamente connessa con il sistema del 'cosa'); b) la mislocalizzazione spaziale corporea legata alle interazioni di spazio e tempo, per esempio l'illusione 'tau' (Helson e King, 1931) in cui tre diverse stimolazioni tattili sull'avambraccio in ordine sequenziale cambiano in base all'intervallo temporale (se l'intervallo tra i due punti stimolati è breve rispetto al contatto tra il secondo e terzo punto, la distanza tra il primo e secondo contatto viene percepita più corta rispetto alla distanza tra secondo e terzo contatto). Un altro esempio di mislocalizzazione è un movimento illusorio chiamato " movimento phi" o "movimento beta". Questa illusione, ben nota in ambito visivo e nota come terza legge del movimento apparente di Korte (1915), viene utilizzata anche con la modalità tattile con dei colpi vibrotattili di 150Hz che 'picchiettano' la pelle in modo periodico costante creando l'effetto di un flusso, piuttosto che di singoli punti (vedere per esempio, Sherrick e Rogers, 1966). Un'altra dislocazione viene chiamata "illusione del coniglio" o "illusione del salto", proprio perché si ha la percezione come se un coniglio salisse su per l'avambraccio. Anche per questa illusione colpetti distribuiti equamente in modo spaziale e temporale vengono percepiti come un movimento continuo (si verifica, per esempio, se vengono dati con una successione temporale costante 15 piccoli colpetti in tre siti di contatto equidistanti nell'avambraccio, 5 colpetti per ciascun sito) (Geldard e Sherrick, 1972; Flach e Haggard, 2006). c) Il fallimento nei cambiamenti in un pattern spaziale sulla pelle, per esempio il caso riportato da Gallace e coll. (2006) in cui vi è un'incapacità a percepire cambiamenti spaziali (elementi in più o in meno) per semplici pattern/campioni tattili. Questi eventi tattili spaziali sono presentati in sequenza insieme a un *mask* tattile ("*mudsplash*" tattile) che coincide con il cambiamento spaziale (analogo al *change blindness:*incapacità a rilevare grandi cambiamenti in una scena visiva o uditiva).

Localizzazione spaziale esterna al corpo, riferita all'esplorazione aptica attiva e definita come l'abilità di localizzare dei punti nello spazio esterno al corpo. La letteratura in merito riporta alcune distorsioni, concernenti soprattutto la natura dell'esplorazione, per esempio sulla capacità dei soggetti di ritornare in una posizione precedentemente toccata e la loro abilità nel riportare dove questa localizzazione sia nello spazio (Klatzky, R. L. & Lederman, S. J., 2003). Da notare che mentre per la prima domanda la performance può dipendere da memoria motoria, per la seconda la prestazione dipende dalla rappresentazione spaziale costruita sull'elaborazione aptica.

La maggior parte degli errori sistematici riguardano i punti di riferimento che assumono i soggetti, egocentrico o allocentrico. Per il tatto, i punti di riferimento cambiano in base al compito e alla postura del soggetto. Shimono e coll. (2001) hanno riscontrato che per il punto di riferimento egocentrico cambiava in base alla distanza degli oggetti e alla mano usata (veniva chiesto ai soggetti di allineare degli oggetti e diverse distanze dal loro corpo e gli oggetti dovevano puntare il soggetto). Per quanto riguarda il punto di riferimento allocentrico è possibile localizzare gli oggetti in base agli assi intrinseco (per esempio i bordi di un tavolo) o utilizzare un sottoinsieme di oggetti per definire un punto di riferimento rispetto agli altri. A volte si riscontra che in compiti che richiedono una localizzazione allocentrica possono essere risolti facendo riferimento a più di un sistema di riferimento. Ne è un esempio lo studio di Kappers e coll. (2007) in cui gli autori chiedono l'orientamento di barre relativamente una con l'altra ed in particolare di aggiustare la direzione di una barra in modo tale che sia parallela ad una barra localizzata sullo stesso piano. Gli autori analizzando i tipi di errori concludono che i soggetti hanno usato due tipi sistemi di riferimento contemporaneamente, uno centrato sul corpo (probabilmente la mano) e l'altro considerando lo spazio esterno (Kappers e coll., 2003). Gli autori aggiungono che vi è una differenza interindividuale dei soggetti nel preferire (o usare di più) un punto di riferimento rispetto all'altro.

Una questione fondamentale riguardo all'esplorazione spaziale aptica riguarda il fatto se la distanza metrica sia uniforme rispetto allo spazio, indipendentemente dalla grandezza della distanza e direzione, il che costituirebbe un'isotropia. Sembra invece che la maggior parte delle illusioni del senso aptico (es. l'illusione radiale/tangenziale, in cui estensioni lineari sono percepite come più lunghe se radiali anziché tangenziali, Cheng 1968 e Marchetti e Ederman, 1983; e verticale/orizzontale, ovvero la percezione tattile di T o L in base al piano: se piano orizzontale le linee verticali sono sovrastimate rispetto alla lunghezza delle componenti orizzontali, Burtt, 1917) dipenda proprio dal fatto che sia anisotropo. La grandezza dell'effetto dell'anisotropia sembra dipendere dal tipo di movimento utilizzato per esplorare la grandezza (Heller e coll. 1997; Wong e coll., 1977).

In merito all'integrazione del **'cosa'** e del **'dove'**, partendo dalle diverse proprietà degli stimoli aptici, materiali e geometriche, Purdy, Lederman e Klatzky (2004) hanno ipotizzato due diversi modelli di elaborazione dell'identità e della localizzazione. Secondo gli autori solo l'identificazione delle proprietà geometriche comporta dei costi, mentre ciò non si verifica per l'identificazione delle proprietà materiali. Gli autori ipotizzano che queste differenze non siano solo in termini di elaborazione ma anche una differenza relativa alla localizzazione ipotizzando

due modelli di elaborazione corrispondenti: *"the Intercept Cost model"* e *"the Slope & Intercept Cost Model*. L'analisi delle pendente e intercetta nei due modelli, a livello matematico, rileva il semplice fatto che l'elaborazione della localizzazione in un compito di identificazione può incrementare il tempo di elaborazione. Il primo modello predice che aggiungendo l'elaborazione della localizzazione, questo si ripercuote nel valore dell'intercetta di una funzione (l'elaborazione della localizzazione rappresenta un evento che può accadere e questo sarà rappresentato come valore dell'intercetta della funzione, piuttosto che nel valore dell'inclinazione), mentre il secondo modello predice un incremento sia del valore dell'intercetta che l'inclinazione. In altre parole, dall'analisi di entrambi i modelli implicati in diversi compiti, gli autori concludono che vi è un costo legato alla richiesta di localizzare un oggetto conosciuto, e la localizzazione non è elaborata automaticamente. Gli autori ipotizzano che la richiesta della localizzazione implica una risposta con un gran numero di alternative (rispetto ad un compito di identificazione), inoltre visto che nel tipo di compito svolto veniva chiesta la localizzazione indicando il nome del dito, gli autori indicano anche questo fattore come influente per la temporalità della risposta (richiede un tempo aggiuntivo di elaborazione) ed infine suppongo anche che se la localizzazione non viene esplicitamente richiesta dal compito, non viene elaborata automaticamente.

Altri studi sull'integrazione delle caratteristiche del senso aptico hanno analizzato gli effetti di interferenza del "*what*" vs. "*where*" nei compiti percettivi. Per esempio, Purdy, Lederman e Klatzky (2004) somministrano un compito di detezione di caratteristiche per stimoli tattili presentati in diverse dita di ciascuna mano. I risultati mostrano che i tempi di reazione erano più lunghi quando erano richiesti e non solo quando dovevano richiamare l'informazione sulle caratteristiche, ma anche riportare quale dito fosse stato stimolato, rispetto alla condizione in cui il richiamo spaziale non era richiesta. Gli autori concludono che la localizzazione tattile non viene elaborata automaticamente per le caratteristiche degli oggetti. Questo studio riguarda l'influenza della codifica spaziale sulla detezione dell'identità dello stimolo tattile, ma non dice nulla sull'influenza dell'identità dello stimolo su un compito spaziale.

Chan e Newell (2008) mostrano questa distinzione con uno studio comportamentale per un compito-dipendente, mostrando una distinzione cosa/dove che trascende le modalità, usando un paradigma del doppio compito. Compiti simultanei di 'what' and 'where' sembrano interferire mutualmente più in compiti *crossfunction* in entrambe le condizioni cross modali e intramodali, indicando un set di risorse che dipende dalla richiesta del compito e non dalla modalità utilizzata per eseguirlo. Dopo il primo esperimento, affermano che l'elaborazione del

'cosa' e del 'dove' sia dissociabile nel dominio aptico. Essi usano un paradigma di interferenza del doppio compito in cui ai partecipanti veniva richiesto di svolgere un compito primario di riconoscimento e contemporaneamente dovevano svolgere un compito di interferenza. In una condizione entrambi i compiti richiedevano il dominio del 'cosa' (es. Forma e rugosità) e del 'dove' (es. localizzazione e orientamento), in un'altra condizione i due compiti richiedevano domini diversi. I risultati mostrano che quando è presente una intereferenza tra i due compiti (stesso dominio) i partecipanti hanno difficoltà nel svolgere il primo compito, sia nel dominio del 'cosa' che del 'dove'. Gli autori concludono dicendo che l'informazione viene elaborata indipendentemente per il riconoscimento e per la localizzazione spaziale.

Entrambi gli studi (Purdy e coll., 2004 e Chan e Newell, 2008) mostrano che vi sia un'indipendenza dell'elaborazione dell'informazione spaziale e identità. Questa dissociabilità tra l'identà e la localizazzione sembra contrastare con le evidenze sull'automaticità del *feature-to-location binding* per la modalità visiva e uditiva.

Delogu e coll. (2013) reputano che questa differenza tra le evidenze del dominio tattile vs visivo e/o uditivo, dipende dal fatto che il disegno sperimentale in questi studi è diverso rispetto agli studi sul binding e memoria di lavoro. Ragion per cui, conducono due esperimenti che hanno come obiettivo il verificare se la tessitura e la localizzazione di oggetti esplorati in modo aptico sono mantenuti in WM in modo integrato o indipendente, svolgendo compiti in una condizione in cui entrambe sono rilevanti per il compito, e in una condizione in cui solo la posizione o solo la trama devono essere mantenuti in memoria per il richiamo futuro. Gli autori riscontrano che anche per la modalità aptica, esiste un biding del ‘cosa’ e del ‘dove’, concludendo che l’attenzione all’associazione tra le caratteristiche non è necessaria per il *feature-to-location binding* nella memoria di lavoro aptica, ma che certamente, per essere integrati occorre che il soggetto codifichi e mantenga in memoria entrambe.

In conclusione quindi, sebbene la maggior parte degli studi sono a favore per una indipendenza tra i due tipi di elaborazione, a seconda del tipo di compito effettuato, e del processo sottostante specifico, possono essere integrati.

***Binding* ‘*what*’ and ‘*where*’ nel sistema aptico: network neurali e finestra temporale.**

Studi anatomici e fisiologici del sistema somatosensoriale generalmente supportano un modello seriale e gerarchico tra le cortecce primaria e secondaria (es. Mishkin, 1979; Iwamura, 1998).

Diversi studi riportano che l’informazione somatosensoriale viene trasmessa tramite circuiti gerarchici (Hyvärinen and Poranen 1978; Iwamura and Tanaka 1978; James e coll. 2007),

partendo da circuiti talamocorticali che trasmettono l'informazione nella corteccia somatosensoriale primaria (SI) localizzata nel giro postcentrale del lobo parietale (DiCarlo e coll. 1998; Huffman and Krubitzer, 2001) e che i neuroni in SI la trasmettono a loro volta in aree somatosensoriali di livello superiore, come la corteccia somatosensoriale secondaria (SII) e il lobo parietale superiore (SPL), che elabora rappresentazioni di forme più complesse (Roland et al. 1998). Affinché vi sia il riconoscimento dell'oggetto, queste aree sono connesse con la corteccia insulare, regioni prefrontali remote e la corteccia cingolata anteriore (ACC), pensato come circuito fronto-parietale implicato nel riconoscimento di oggetti complessi tramite il senso aptico (Binkofski e coll. 1999; Stoeckel e coll. 2003).

Sebbene non riportati in dettaglio in questa sede, bisogna tener presente che diversi studi con pazienti con lesioni cerebrali hanno corroborato la divisione delle due vie del 'cosa' e del 'dove' a livello somatosensoriale (es. Caselli 1991, 1993) e sembra che regioni parietali sono associate con il disturbo selettivo nella memoria di forme geometriche (studi su lesioni con conseguente agnosia tattile, Caselli 1991; Reed e coll. 1996).

Riguardo al riconoscimento di oggetti e apprendimento tattile (memoria di forme), diversi studi di neuroimmagine rilevano l'attivazione della via somatosensoriale ventrolaterale (es., Bonda e coll., 1996; Ginsburg e coll., 1987; Roland e coll., 1998). Questi studi, inoltre, corroborano l'ipotesi del coinvolgimento della corteccia SII nell'elaborazione del riconoscimento degli oggetti usando il sistema aptico (es. Caselli, 1001) e il ruolo del lobo parietale (vedi agnosia tattile).

Reed e coll. (2005) mostrano che il riconoscimento aptico degli oggetti e la loro localizzazione attivano le aree parietali inferiori ventrolaterali e superiori dorsali, rispettivamente, suggerendo una correlazione con la distinzione tra le vie ventrale e dorsale del sistema visivo discusse da Ungerleider and Mishkin (1982) in ambito visivo. In particolare, compiti di riconoscimento tattile degli oggetti (ignorandone la localizzazione) attiva il lobo frontale, parietale inferiore bilaterale e regioni prefrontali sinistre (implicate nell'integrazione di caratteristiche e nell'elaborazione del nome specifico, l'identità dell'oggetto). Mentre compiti di localizzazione tattile degli oggetti (ignorandone l'identità) attivano aree parietali superiori bilaterali (implicate nell'elaborazione spaziale).

James e coll. (2007) ipotizzano che il riconoscimento degli oggetti aptici procede secondo due network neurali paralleli, uno per l'analisi delle proprietà geometriche, l'altro per l'analisi delle proprietà materiali. Gli autori riportano diversi studi in cui la sola vibrazione o palpazione di oggetti ha mostrato che il giro postcentrale contiene quattro mappe somatotopiche distinte corrispondenti alle Aree 3a, 3b, 1 e 2 (Bodegard e coll., 2001; Burton e coll., 1997a; Mglone e

coll., 2002), ed anche alcuni studi post-morte che rilevano l'esistenza fi almeno quattro mappe somatotopiche separate in SII (Eickhoff e coll., 2006). Dopo una metanalisi di vari studi, gli autori sostengono che queste due vie si differenziano già negli stadi precoci, nella corteccia SI, in cui l'Area 1 è specializzata per la microgeometria, e l'Area 2 per la macrogeometria (si rifanno agli studi di Roland e coll, 1998 e Servos e coll. 2001 in cui le proprietà materiali utilizzate erano la tessitura e la durezza) . Aggiungono che il sistema aptico è organizzato in stadi di elaborazione gerarchici e che ogni stadio è rappresentato da una regione cerebrale specifica. La corteccia SII per esempio, elabora le caratteristiche di alto livello delle proprietà materiali; mentre l'aspetto anteriore del solco intraparietale (aIPS) e l'area laterale occipitale complessa (LOtv) (James e coll., 2002; Peltier e coll., 2007; Prather e coll., 2004; Stoesz e coll., 2003; Zhang e coll. 2004) si attivano durante compiti in cui è richiesta l'elaborazione delle proprietà geometriche. Vi sono studi discordanti: per esempio circa la distinzione dell'Area1 e Area2 per l'analisi della micro e macrogeometria (Kitada e coll., 2006; Stoesz e coll., 2003) ed anche la specializzazione dell'aIPS e LOtv (Servos coll., 2001; Stoesz e coll., 2003). Il convolgimento della SII per l'elaborazione cognitiva di alto livello nel riconoscimento degli oggetti tramite il senso aptico sembra avere maggiore consenso dagli studi riportati dagli autori (oltre a quelli citati, anche Bonda e coll., 1996; Pietrini e coll., 2004; Reed e coll. 2004). Diversi studi riportati dagli autori (Binkofski e coll., 1999; 1998) mostrano che la regione aIPS non sia puramente somatosensoriale, ma che integri informazioni somatosensoriali e motorie.

Bud e Timora (2013) riportano che nella modalità somatosensoriale l'ampiezza modulata dalla stimolazione vibratoria elicita la risposta SSR massima tra i 20 e i 30 Hz (Snyder, 1992; Tobimatsu et al., 1999; Langdon et al., 2011). In particolare, nello studio di Snyder (1992) gli SSR sono stati registrati in risposta ad uno stimolo vibratorio sulle dita o sul palmo della mano. Egli riscontra il picco del segnale EEG a circa 26Hz e la latenza degli SSR era a 58±14 msec. Nello studio di Tobimatsu e coll (1999) lo stimolo vibratorio era anche sul palmo destro e riscontrano il picco di risposta a 21 Hz. Infine, nello studio di Langdon e coll. (2011), essi riscontrano la risposta massima a 30 Hz.

Gurtubay-Antolin e coll. (2015) indagano la velocità dell'elaborazione aptica tramite registrazione ERPs, durante un compito di discriminazione delle forme (senza informazione visiva). In particolare, dopo una breve esplorazione, i partecipanti devono giudicare se l'oggetto toccato corrisponde ad un oggetto atteso (nome presentato sullo schermo). Gli autori riscontrano per ogni incongruenza una negatività frontocentrale a 175 ms. Inoltre, dopo un'analisi della sorgente, gli autori riscontrano che questa attività differenziale è localizzata nelle aree somatosensoriali di livello superiore e nelle regioni prefrontali implicate nel

monitoraggio degli errori e nel controllo cognitivo. Gli autori interpretano questi risultati come prova che il sistema somatosensoriale riesca a completare l'elaborazione aptica sostanziale affinché si rilevi una risposta contrastante nelle cortecce mediali e prefrontali in meno di 200 ms e che il sistema aptico ha un riconoscimento veloce e connesso con il monitoraggio degli errori e conflitto. Questo dato è in linea con gli esperimenti comportamentali menzionati di Klatzky and Lederman (1985; 1995).

1.1.3 Cosa sappiamo dell'udito: 'cosa' e 'dove' nell'udito

Definizione e aspetti generali del *binding* unisensoriale uditivo.

Un primo approccio di analisi del problema della integrazione percettiva uditiva, ed in particolare della distinzione dei suoni in ambienti complessi, è stato presentato da Bregman nel 1990. Egli coniò l'espressione ***Auditory Scene Analysis* (ASA)** per indicare una situazione comune di ascolto dell'esperienza quotidiana che include tre aspetti fondamentali della percezione acustica: segmentazione, integrazione e segregazione. Con questa espressione (per altro presa dalla '*scene analysis*' utilizzata nella ricerca visiva) l'autore indica il processo per cui più sorgenti ambientali vengono percepite come un'unità percettiva o divise. Egli si chiede infatti come gli ascoltatori siano capaci di identificare, localizzare e caratterizzare individualmente le sorgenti sonore, per esempio in ambienti dove sono presenti più suoni. Il problema del *binding* unisensoriale applicato al sistema uditivo, si preoccupa proprio di esplorare questi interrogativi.

Il termine "scene analysis" è stato coniato in ambito visivo, per lo studio delle strategie usate dal computer per unificare o meno proprietà visive di oggetti diversi o degli stessi oggetti (es.texture, colori, distanza, ecc-). Per analogia, l'ASA è il processo attraverso cui gli elementi uditivi di una scena sono percepiti come unità percettive.

L'ASA indica, quindi, sia il problema che il processo percettivo (Bregman, 2008) e apre il problema di come il sistema uditivo organizzi i suoni provenienti da diverse fonti nell'ambiente, e quindi come poter distinguere un suono discreto o una sequenza di suoni. Ci si interroga fondamentalmente quindi, su due processi fondamentali: come un individuo segrega e raggruppa le varie componenti dei suoni in un flusso o percetto unico coerente (Bregman, 1990). Bregman (1990) propone l'esistenza di due meccanismi base: 1) meccanismi primari, che elaborano i suoni mischiati in modo automatico utilizzando semplici trasformazioni, 2) meccanismi di schema di base, che dipendono da processi attentivi, intenzionali e di conoscenza pregressa.

Come accennato, l'analisi dell'ASA , intesa come una moltitudine di stimoli acustici che ci circondano, apre il problema dell'identificazione e localizzazione dell'oggetto uditivo. Come per l'oggetto visivo, l'oggetto uditivo può essere considerato come il risultato computazionale della capacità del sistema uditivo di detezione, estrazione e segregazione di regolarità spettro temporali all'interno di un ambiente (appunto, l'ASA) (Bizley e Cohen, 2013). Come per qualsiasi rappresentazione di un oggetto, Bizley e Cohen, 2013 indicano dei processi di rappresentazione: dal dato astratto si estraggono delle regolarità, generalmente correlate con l'importanza dell'informazione semantica al fine di poter svolgere delle decisioni (categorizzazioni) che determinano il comportamento. Gli autori definiscono l'oggetto uditivo come un costrutto percettivo, corrispondente ad un suono ed ad una determinata sorgente. Si tenga conto che anche una sequenza di oggetti può determinare un oggetto di per sé. Gli autori indicano alcune caratteristiche dell'oggetto uditivo: 1) è una conseguenza di azioni o eventi, determinato da una intenzionalità di fondo (es. linguaggio); 2) è possibile estrarre le sue costituenti e proprietà spettro temporali che lo distinguono dagli altri oggetti uditivi; 3) può essere descritto come il risultato delle sue caratteristiche di altezza, timbro, forza, armonicità, differenza binaurale ecc; 4) è invariante rispetto al contesto (una nota di violino rimarrà sempre una nota di violino anche in diversi contesti); 5) è possibile generalizzarlo e prevederlo (es. immaginare il suono anche se non è presente).

Dopo aver capito che le sorgenti sonore non forniscono una chiara decomposizione del segnale, entra in gioco il problema su come avvenga l'ASA in contesti quotidiani, complessi. Secondo Bregman sono necessari tre processi per l'analisi e decomposizione delle sorgenti sonore. I primi due riguardano aspetti già appresi dall'ascolto: l'automaticità (per esempio, immaginare di ascoltare il proprio nome in una scena caotica: l'ipersensibilità e l'automaticità con cui viene percepito il proprio nome, che rende il proprio nome particolarmente 'presente') e la volontarietà (per esempio, quando ci si attende che qualcuno ci chiama). Oltre a questi due processi ne è necessario un terzo, non direttamente legato all'ascolto precedente, che Bregman riferisce ad una classe di 'metodi generali' definiti con l'espressione *'primitive auditory scene analysis'* (Bregman, 1990, pag.38), ovvero dei meccanismi che analizzano delle proprietà acustiche generali che possono essere usate per qualsiasi sorgente e qualsiasi contesto. Queste proprietà generali della produzione delle sorgenti sonore non sono specifiche per nessun tipo di classi di suoni, ma possono essere, per esempio, inferite da regolarità acustiche (per esempio, la maggior parte dei suoni sono caratterizzati da armoniche, ovvero da una frequenza fondamentale e dalle frequenze componenti (multipli della frequenza fondamentale). Bregman (1993) elenca alcune delle regolarità:

1-suoni con correlati raramente iniziano o finiscono esattamente allo stesso tempo;
2-la gradualità del cambiamento: a)un singolo suono tende a cambiare le sue caratteristiche regolarmente (facilmente) e lentamente; b)una sequenza di suoni provenienti dalla stessa sorgente tende a cambiare le sue proprietà lentamente;
3-quando un corpo vibra con un tempo ripetitivo, le sue vibrazioni danno vita ad un pattern/schema acustico nel quale le frequenze componenti sono multiple di una fondamentale comune;
4-molti cambiamenti che avvengono in un evento acustico influenzeranno tutte le componenti del suono risultante nello stesso modo e nello stesso tempo.
Bregman con la sua teorizzazione, in qualche modo ha compreso ampi campi di ricerca tra cui psicoacustica, percezione linguistica, teoria della musica e composizione e *computer modeling*.
Nel 1996, Daniel P. W. Ellis propone un modello computazionale, nonchè modello alternativo alla ASA di Bregman, denominato "*the prediction-driven approach*". In questo modello l'analisi è un processo di riconciliazione tra le caratteristiche acustiche osservate e le predizioni di un modello interno dei suoni presenti nell'ambiente. In questo modo, i suoni predetti fanno parte dell'interpretazione della scena, tanto più coincidono con l'input sonoro. Ellis quindi, non fa altro che dare importanza ai processi top-down nella costruzione della scena ambientale.
Darwin C J, (1997) indica i principali tipi di raggruppamento uditivo:il raggruppamento sequenziale e il raggruppamento simultaneo. Come riporta l'autore già i compositori barocchi si divertivano a comporre dello polifonie, implicitamente legate ai principi di raggruppamento sequenziale (per esempio nel trillo). Un esempio dell'importanza di questo tipo di raggruppamento, o meglio, il suo fallimento, è dimostrata dall'illusione di Wessel (Wessel, 1979) in cui un'esecuzione rapida di tre note che aumentano in altezza, ma che sono alternate di timbro, creano un effetto diverso se suonate lentamente o velocemente, ovvero nel momento in cui si eliminano le armoniche più basse dalle note dispari si ha un timbro brillante, mentre se si tolgono le armoniche più alte dei suoni pari si ha un timbro spento. Darwin e Bethell-Fox (1977) creano un effetto di flusso tramite la modulazione della frequenza principale (altezza del suono). Se infatti viene un ripetuto un determinato schema (per esempio delle sillabe ripetute 'yaya') alternando due frequenze diverse, dopo un paio di ripetizioni si sentirà come se si sdoppiasse in due flussi, uno dei quali è 'silenzioso' e l'altro risulta in uno schema modificato (per esempio le sillabe 'gagagaga'). Un effetto simile a quello di Wessel (1979) è anche dato dai cambiamenti della localizzazione della fonte sonora, per esempio alternando i suoni dello stimolo di Wessel dall'orecchio destro e sinistro (piuttosto che cambiando timbro) (Deutsch, 1979). In altre parole, Darwin afferma che la continuità del timbro, dell'altezza e

della localizzazione spaziale, sono importanti per mantenere l'integrità percettiva della sorgente sonora nel tempo (Darwin, 1997). Per quanto riguarda il raggruppamento simultaneo, l'autore porta l'esempio dell'abilità di ascoltare due strumenti distinti in simultanea, e questa abilità viene meno nel momento in cui i suoni prodotti sono gli stessi e presentano stessa frequenza o di rapporto simile (come per esempio l'ottava di quinta). Darwin riporta due esempi legati all'altezza e all'asincronia, come fondamentali per la separazione di suoni simultanei. In particolare, Darwin si concentra sulla capacità umana di utilizzare l'armonicità per distinguere due fonti sonore simultanee tramite la risoluzione della frequenza (più facile se vi sono meno armoniche). In particolare, si è interrogato sull'influenza della differenza temporale interaurale (ITDs) che sebbene sia dominante per la localizzazione di suoni complessi, non può essere usata da sola per segregare il linguaggio di un singolo ascoltatore da suoni simultanei simili. ITD riesce a tracciare la fonte di un suono nel tempo. L'autore interpreta questi risultati come indice dell'esistenza di due distinti livelli di elaborazione uditiva, che indicano a loro volta che gli ascoltatori raggruppano in oggetti uditivi piuttosto che usare le frequenze che hanno un comune ITD.

Griffiths e Warren (2004) indicano quattro principi teorici per inquadrare un oggetto uditivo, dove per oggetto uditivo intendono sia l'analisi dell'oggetto che l'oggetto in sé. Gli autori riportano le fasi di analisi dell'oggetto uditivo in un unico modello, partendo da più modelli di riferimento.

1) Secondo il primo principio l'analisi dell'oggetto uditivo può essere basata sull'analisi della sorgente del suono. In ambito uditivo, la distinzione tra la sorgente e l'evento sebbene sia chiara in linea teorica, a livello rappresentazionale lo è meno, nonché la capacità di estrarre le invarianti e trasformazionali. Secondo gli autori, come termine ultimo, l'oggetto uditivo può essere categorizzato in più di un modo, implicando anche una conoscenza a priori.

2) Secondo il secondo principio l'oggetto uditivo dovrebbe essere separato dalle *boundaries* o limiti percettive. Il suono può essere rappresentato in termini temporali o spettrali (frequenze), ragion per cui una rappresentazione bidimensionale frequenza-tempo di un suono complesso non è disponibile se si fa riferimento ad una sola dimensione (ecco il limite percettivo dell'oggetto uditivo). Se si riprende la definizione dell'oggetto secondo Kubovy e van Valkenburg (2001) , che enfatizzano il ruolo dei margini anche in oggetti definiti da frequenza e tempo, il problema si sposta sulla definizione delle diverse frequenze e tempo implicati in un unico suono coerente. Il problema è stato risolto raggruppando insieme delle frequenze o dei tempi, portando a studiare il suono in termini di armoniche o di *onset synchrony*. I suoni in

questo modo sono una combinazione di oggetti. Ne è un esempio il timbro, che riesce a distinguere due suoni nonostante essi abbiano altezza, durata, intensità e localizzazione

3) Secondo il terzo principio l'analisi dell'oggetto uditivo implica l'astrazione dell'informazione che è indipendente dalla specifica rappresentazione sensoriale. In altre parole, l'estrazione di caratteristiche invarianti porta alla costanza dell'oggetto.

In sintesi, quindi, secondo gli autori l'analisi dell'oggetto uditivo include l'analisi dell'informazione (che corrisponde all'oggetto in sé), la separazione dell'informazione dell'oggetto rispetto al resto del mondo sensoriale, la capacità di astrazione dell'informazione (generalizzazione) sia all'interno della stessa modalità sensoriale che tra i vari sensi (quest'ultimo punto, nonché quarto principio, non viene riportato perché oggetto di discussione).

Binding '*what*' and '*where*' nell'oggetto uditivo: studi comportamentali.

Per quanto riguarda il '**cosa**' dell'oggetto uditivo, come già accennato, due dei principi fondamentali nel raggruppamento di caratteristiche per creare un oggetto, sono i principi di raggruppamento sequenziale e simultaneo. Un esempio di uno dei paradigmi comunemente usato per analizzare il raggruppamento sequenziale e la segregazione del flusso uditivo è il paradigma del flusso ABA (Bizley e Cohen, 2013). In questo paradigma due toni suonano a diverse frequenze (A e B) con un intervallo frapposto. Questo schema si ripete in modo sequenziale. Se la differenza tonale tra A e B è bassa (di mezzo semitono), l'ascoltatore solitamente riporta di percepire un unico flusso uditivo; maggiore è la distanza tonale (diversi toni) l'ascoltatore riporterà di sentire due flussi uditivi differenti. In una fase intermedia (intervallo medio tra A e B) l'ascoltatore può sentire uno o due flussi uditivi alternati nelle prove (definito stimolo 'bistabile', dal momento che può essere riportata una delle due alternative possibili). Il paradigma può essere variato manipolando la durata dei toni, la durata dell'ascolto complessivo, il numero di ripetizioni, ecc. (Shamma e Micheyl, 2010). L'intervallo temporale tra A e B sembra essere molto rilevante per la distinzione tra il numero di flussi percepiti riportati (si veda l'ipotesi di coerenza temporale di Elhilali et al., 2009), così come il tempo di A o di B, ovvero di ciascun tono (Micheyl et al, 2013).

Tra e caratteristiche dell'oggetto uditivo più studiate sono l'altezza e il timbro, di cui si parlerà nella parte neurale.

Christison-Lagay e coll. (2015) riportano che per identificare un oggetto uditivo occorre estrarne delle regolarità, non solo di raggruppamento spettrale, ma anche di regolarità spaziali

e temporali, enfatizzando il fatto che i flussi sonori di per sé variano e fluttuano, e che dobbiamo fare riferimento a ciò che ci consente di generalizzare il riconoscimento di un suono o flusso sonoro.
L'autore far riferimento in particolar modo al fatto che il suono si sviluppi temporalmente e spazialmente e che per esempio, solo grazie alla temporalità creiamo un'aspettativa circa il suono che verrà dopo, per esempio, in un flusso ripetitivo (es. paradigma ABA). L'autore riprende il fatto che la percezione degli stimoli tende a fluttuare spontaneamente nel tempo (si vedano i modelli di rappresentazione della regolarità predittiva, *predictive-regularity representations,* di Denham e Winkler, 2006; Winkler et al., 2009). Aggiunge che bisogna considerare che in realtà un unico suono è l'integrazione di più suoni (inteso come unione di armoniche) e che possiamo unire in un unico oggetto uditivo distinte regolarità di frequenza (van Noorden, 1975), localizzazione spaziale (Hill et al., 2011), timbro (Singh, 1987) e quant'altro (Bregman, 1990; Grimault et al., 2002; Vliegen and Oxenham, 1999).

Per quanto riguarda il sistema del '**dove**', si deve tenere presente che una delle differenze sostanziali tra la modalità visiva e quella uditiva è legata all'elaborazione della spazialità degli eventi. Nella modalità visiva infatti l'informazione spaziale viene elaborata in parallelo, mentre nella modalità uditiva manca la corrispondenza tra la localizzazione spaziale e un campo recettivo sensoriale ad hoc (Ahveninen e coll., 2014).
Studi di neuroimmagine hanno deputato il sistema del "dove" nell'ambito uditivo, in aree uditive posteriori non primarie, includendo il planum temporale (PT) e il giro temporale superiore posteriore (STG), che vengono attivati dei cambiamenti di direzione del suono orizzontalmente, cambiamenti di distanza e di movimento. Come abbiamo visto, però, queste aree vengono anche attivate da altre caratteristiche sonore (es. il timbro), ragion per cui bisogna fare cautela nel pensare a queste aree come puramente 'spaziali'.
La localizzazione sonora spaziale ha importanza per due ragioni: 1) per localizzare le fonti sonore; 2) per separare dei suoni dagli altri in base alla loro posizione spaziale (Blauert, 1997). Senza considerare altri tipi di informazioni aggiuntive: per esempio ascoltando una persona che bussa alla porta, in base alla provenienza del suono possiamo farci un'idea dell'altezza della persona (Tajadura-Jiménez, 2015). Riguardo al secondo punto, se ne è accennato circa l'importanza dell'analisi dell'ASA (Bregman, 1990) e più in generale della distinzione dell'oggetto uditivo da un flusso di suoni e dal rumore di sottofondo (Moore, 1997).
La localizzazione uditiva si basa sull'informazione monoaurale e binaurale (Yost e Gourevitch, 1987). Come ben noto, le due informazioni binaurali sono diverse in tempi di arrivo, il

cosiddetto tempo di differenza interaurale, ITD, o differenza di fase interaurale, IPD (Middelebrooks e Green, 1991) (la differenza tra l'usare l'ITD per basse frequenze e l'IPD per alte frequenze fa riferimento alla teoria doppia di Strutt, 1907). Tra i segnali di localizzazione monoaurale la più importante è il cambiamento della grandezza dello spettro sonoro causato dall'interazione del suono, con la testa il corpo e il padiglione (Blauert, 1997; Macpherson e Sabin, 2007; Middlebrooks e Green, 1991; Shaw, 1966; Wightman e Kistler, 1989). Un altro segnale monoaurale importante è il rapporto di energia diretta a riverberante (direct-to-reverberant energy ratio, DRR) che in altre parole esprime la quantità di energia sonora che raggiunge le orecchie direttamente dalla sorgente contro la quantità che è riverberata dalle pareti degli spazi chiusi (Larsen et al., 2008). Generalmente i segnali monoaurali sono più ambigui di quelli binaurali forse perché il sistema uditivo fa assunzioni a priori sulle caratteristiche sonore per stimare gli effetti di filtraggio dei segnali spaziali monoaurali (Ahveninen e coll., 2014).

La maggior parte degli studi sulla localizzazione della sorgente sonora utilizzano un sistema di riferimento di coordinate sferiche bipolari, prendendo le due orecchie come i due poli e l'origine al centro tra le due orecchie (Duda, 1997). In questo sistema di coordinate l'azimuth (o localizzazione orizzontale) di un oggetto è definita dall'angolo tra la sorgente e l'asse interaurale e la distanza è misurata dal centro della testa. Per determinare la localizzazione sonora in un determinato punto dell'ambiente si possono analizzare le funzioni di trasferimento correlate con la testa (*head-related transfer functions,* HRTFs) e le risposte degli impulsi binaurali della stanza (*binaural room impulse responses,* BRIRs). Queste due risposte o funzioni possono secondo Ahveninen e coll. (2014) forniscono una caratterizzazione acustica completa dell'informazione spaziale, possono variare di poco da ascoltatore in ascoltatore e possono essere usati come indicatori anche in ambienti virtuali.

Come accennato, negli ambienti circoscritti, il riverbero altera lo spettro monoaurale sia dei segnali ITD che IPD. Recenti studi hanno mostrato che in alcuni casi il riverbero può aiutare la localizzazione sonora dal momento che fornisce un'informazione ulteriore al DRR (Bronkhorst e Houtgast, 1999; Kolarik et al., 2013; Kopco e Shinn-Cunningham, 2011; Ronsse e Wang, 2012; Zahorik et al., 2005).

Si è visto che l'informazione binaurale risulta particolarmente importante per la localizzazione di suoni soprattutto quando più suoni sono presentati simultaneamente (es. con stimoli linguistici: Drullman e Bronkhorst, 2000; Hawley et al., 1999; Simpson et al., 2006). Diversi studi mostrano che le abilità di localizzazione dipendono sia da risorse attentive (attenzione

spaziale selettiva diretta), sia della conoscenza a priori dell'ambiente che dal contesto (distanza spaziale e temporale delle altre sorgenti sonore non rilevanti).

Delogu, Nijboer, e Postma (2012) indagano le interazioni tra le informazioni spaziale e temporale nella memoria di lavoro uditiva (*binding 'where'* e *'when'*). Gli autori svolgono due esperimenti in cui ai partecipanti vengono presentate delle sequenze di suoni provenienti da diverse posizioni dello spazio e viene chiesto di ricordare sua la posizione che l'ordine seriale. Nel primo esperimento l'attenzione durante la codifica viene manipolata contrastando blocchi 'puri' (trials solo di localizzazione o solo di ordine seriale) a blocchi 'misti' (con diversa percentuale di trials di tipo spaziale o solo di ordine seriale). Nel secondo esperimento, i blocchi 'puri' sono contrastati dai blocchi in cui le prove spaziali e di ordine seriale sono mescolati con un terzo compito che richiede una categorizzazione semantica dei suoni. I risultati di entrambi gli esperimenti mostrano che, mentre il richiamo di ordine seriale è inficiato linearmente dalla codifica simultanea di una caratteristica concorrente, il richiamo della posizione non è inficiato dalla concomitante codifica della caratteristica. In modo contrastante, il livello della performance globale è più basso per il richiamo spaziale rispetto al richiamo seriale. Gli autori concludono che gli items seriali e di localizzazione sembrano essere codificati indipendentemente nella memoria di lavoro. L'ordine seriale è più facile da richiamare ma fortemente influenzato dall'elaborazione delle concomitanti dimensioni dell'item, mentre la localizzazione dell'item è più difficile da richiamare ma relativamente automatica, come mostrato dalla forte resistenza ad interferire le dimensioni nella codifica/ alle dimensioni interferenti nella codifica.

Delogu e coll. (2014) distinguono due aspetti del binding "*what*" e "*where"* nella memoria di lavoro uditiva. Metodo: compito di riconoscimento (suoni variano per timbro, tono e localizzazione). Gli autori concludono che le caratteristiche di identificazione dei suoni vengono elaborate anche quando non richieste dal compito, mentre la sorgente del suono (localizzazione) non viene considerata se non è richiesta.

***Binding* '*what*' and '*where*' nell'oggetto uditivo: network neurali e finestra temporale**

Studi di neuroimmagine riportano attivazioni corrispondenti alla percezione degli ascoltatori. Sembra essere rilevante quindi, la predittività di un suono o di uno schema di suoni, così come le 'alterazioni' rispetto a questa prevedibilità dello schema sonoro (si veda il modello di predittività regolare di Winkler et al., 2009). Per esempio nello studio di Brendixen e coll. (2009) la registrazione degli ERP mostrava una risposta neurale anche quando un suono non veniva presentato ma era "previsto" nella sequenza di suoni.

Come per l'elaborazione visiva, in parallelo, anche per quella uditiva è stata proposta l'esistenza di due vie di analisi parallela del 'dove' e del 'cosa', rispettivamente via dorsale e ventrale. Anche l'elaborazione gerarchica come possibile ipotesi: da aree di elaborazione precoce ad aree di elaborazione complessa.

A livello corticale, una prima distinzione tra queste due vie viene riportata per la prima volta in primati non umani, non riportati in questa sede. Più di recente, studi di neuroimmagine hanno riportato pattern neurali distinti anche gli umani (studi con soggetti normodotati e pazienti con lesioni cerebrali). Alcuni studi di neuroimmagine hanno corroborato l'idea di una distinzione delle vie della corteccia non primaria del 'cosa', anteriore, e del 'dove', posteriore (Ahveninen et al., 2006; Barrett and Hall, 2006; Rauschecker, 1998; Rauschecker and Tian, 2000; Rauschecker et al., 1995; Tian et al., 2001).

Studi di neuroimmagine hanno dimostrato che compiti spaziali implicano l'attivazione della corteccia uditiva posteriore (dorsale) e che compiti non spaziali coinvolgono il lobo temporale (Alain et al, 2001; Maeder et al, 2001; Arnott et al. 2004).

Per quanto riguarda il sistema del 'cosa' Una delle caratteristiche più studiate nell'identificazione dell'oggetto uditivo è l'elaborazione dell'altezza di un suono. Un'area sensibile all'altezza di un suono è stata identificata nel giro di Heschl (Patterson et al, 2002; Warren et al. 2003) .così come in zone periferiche e centrali dell'elaborazione uditiva. Le ipotesi fatte circa l'elaborazione dell'altezza sonora sono principalmente due: 1) o si tratta di una rappresentazione invariante del planum temporale, che corrobora l'idea di una elaborazione distribuita (e in questo caso il tono è usato come un cue di raggruppamento, Darwin et al., 1997); 2) o ha come area centrale il giro di Heschl, che corrobora l'idea di una elaborazione gerarchica.

Per quanto riguarda il timbro, ci si chiede se richieda una rappresentazione esplicita o implicita. Anche qui, le aree di rappresentazione del timbro sono distribuite sia in zone centrali che nelle regioni di cintura della corteccia uditiva (es. Kumar et al 2007). Ad ogni modo, il timbro sembra una caratteristica modulata da altre caratteristiche sonore, come il tono o la localizzazione spaziale. Sembra che vi sia una prima fase in cui viene analizzato il timbro e viene poi modulato dal tono (Walker et al, 2011) il che secondo Bizley e Cohen (2013) indicherebbe una rappresentazione implicita sua dell'oggetto in se sia della trasformazione dell'oggetto sonoro, per esempio in base alla sua localizzazione spaziale. Studi di neuroimmagine mostrano che neuroni del planum temporale codificano delle rappresentazioni invarianti dello spettro sonoro, una delle chiavi determinanti del timbro (Kumar et al 2007). Un modello di dinamica causale ha identificato un'architettura di elaborazione seriale in cui l'informazione timbrica origina nel

giro di Heschl, viene trasmessa nel planum temporale e nel giro temporale superiore. Secondo questo modello l'estrazione spettrale è completata quando l'informazione raggiunge il planum temporale. Questo modello di elaborazione gerarchica potrebbe spiegare come il timbro ci permetta di identificare uno strumento musicale tra diversi toni e melodie.

Riprendendo il discorso sul raggruppamento sequenziale e simultaneo e il paradigma ABA, Pressnitzer e coll. (2008) riportano che un'attivazione sottostante al paradigma del flusso ABA in stadi precoci nel nucleo cocleare e quindi in zone uditive periferiche.

Per quanto riguarda il sistema del 'dove' Ahveninen e coll. (2014) scrivono una review sulle basi neurali e psicofisiche della localizzazione sonora negli esseri umani. Contrariamente alle altre modalità sensoriali, lo studio della corteccia uditiva è particolarmente arduo, sia per posizione che per diversità (per esempio non sembra esserci un'organizzazione topografica come per il sistema tattile, ma è molto più complesso). In particolare, il candidato migliore per la via del 'dove' è il planum temporale (PT) e il giro temporale superiore posteriore (STG), che presentano una forte attivazione nei cambiamenti di direzione sonora orizzontale (Ahveninen et al., 2006; Brunetti et al., 2005; Deouell et al., 2007; Tata and Ward, 2005), di movimento (Baumgart et al., 1999; Formisano et al., 2003; Krumbholz et al., 2005; Warren et al., 2002), di distanza di segnali indipendentemente dalla intensità (Kopco et al., 2012) e in condizioni in cui è richiesta la separazione di sorgenti sonore multiple (Zündorf et al., 2013). Come accennato, rimane ancora non del tutto chiaro come la modalità uditiva codifica lo spazio acustico: se vi sia un'organizzazione topografica apposita, o i neuroni a livello della corteccia non primaria sono sintonizzati su segnali come l'ITD o l'IDL usando un sistema di codifica a due canali, per non parlare della codifica parallela o gerarchica.

Vi sono pochi studi che hanno studiato l'attivazione subcorticale dei segnali uditivi, sia con tecnica fMRI (Thompson et al., 2006) che con EEG (Junius et al., 2007; Ozmen and Ungan, 2009); mentre diversi studi hanno studiato le vie che contribuiscono al controllo cognitivo di orgine superiore dell'elaborazione spaziale nella corteccia uditiva, incluse le regioni parietali posteriori (es. solco intraparietale) e le regioni frontali (per esempio la corteccia premotoria e corteccia prefrontale laterale).

Alain e coll. (2001) registrano i potenziali evocati (ERPs) durante compiti di discriminazione uditiva di tono e di localizzazione (compito di distinzione di due stimoli S1 e S2, diversi per tono e localizzazione, confrontano sempre se l'S2 rispetto all'S1è identico o no, e se no specificare in che senso è diverso:più alto-basso; destra-sinistra-stesso lato). Rispetto alla localizzazione, l'elaborazione del tono mostra attivazione più grande nella corteccia uditiva e nel giro frontale inferiore. Di contro, la localizzazione genera maggiore attivazione nella

corteccia temporale posteriore, corteccia parietale e solco frontale superiore. Effetti diversi legati al compito sono presenti negli ERPs nelle regioni cerebrali posteriore e anteriore a partire da 300 ms dopo lo stimolo e durano per diverse centinaia di millisecondi. Gli autori concludono circa l'attivazione di due vie diverse per l'identificazione e la localizzazione di stimoli identici,il che corrobora la teoria di una segregazione delle vie del 'cosa' e del 'dove anche per la modalità uditiva (similare alla distinzione fatta per la modalità visiva).

Maeder e coll. (2001) conducono uno studio fMRI come evidenza che riconoscimento di suoni e localizzazione sono elaborati da popolazioni neurali distinte. Soggetti normodotati svolgono 3 compiti: a) confronto di stimoli spaziali simulati con differenze di tempo interaurali, b) identificazione di suoni ambientali, c) riposo. I risultati mostrano che il riconoscimento e la localizzazione di suoni attivano collicolo inferiore, corpo genicolato mediale, giro di Heschl, e parti delle convessità bilaterali temporale, parietale e frontale. Il giro temporale mediale e il precuneo bilaterale e la parte posteriore del giro frontale inferiore sinistro erano più attivati durante il riconoscimento piuttosto che durante la localizzazione. La parte più bassa del lobo parietale inferiore e parti posteriori del giro frontale inferiore e mediale erano più attivati, bilateralmente, per la localizzazione piuttosto che per il riconoscimento. La segregazione delle due diverse regioni di attivazione era più evidente nelle donne. Il paradigma dell'ascolto passivo corrobora l'ipotesi di due vie separate del giro temporale superiore e lobo parietale inferiore, rispettivamente ritenuti coinvolti per il riconoscimento e la localizzazione sonora.

Griffiths, T. D. e Warren, J. D. (2002) svolgono un'analisi del ruolo del planum temporale nella segregazione e appaiamento degli schemi spettro temporali e costruiscono un modello computazionale di questo centro. In particolare, l'informazione spettro temporale conseguente da questo centro è una sorta di cancello per le aree corticali superiori deputate al riconoscimento degli oggetti e la percezione dello spazio uditivo.

Arnott e coll. (2004) svolgono una metanalisi degli studi fMRI e PET per indagare il modello a due vie (dorsale e ventrale) negli umani. Gli autori confrontano studi con compiti di tipo 'spaziale' con studi 'non spaziali'. Gli autori confermano l'ipotesi a due vie riscontrando che compiti di identificazione sonora implicano l'attivazione della via ventrale, mentre compiti di localizzazione sonora attivano aree della via dorsale e aree posteriori della corteccia uditiva. In particolare, il 41% degli studi non spaziali ha riportato un'attivazione del lobo parietale inferiore (nessuno di quelli spaziali). Il 51% degli studi spaziali hanno riportato attività nel solco frontale superiore (anche il 7% degli studi non spaziali riporta l'attivazione in queste aree). Il 56 % degli studi non spaziali riporta un'attivazione dell'area frontale inferiore (aree di Brodmann 45 e 47) (riportata anche in 9% degli studi spaziali). Infine, compiti non spaziali

comportano l'attivazione del lobo temporale in modo distribuito, mentre compiti spaziali riportano attivazione per lo più nella parte posteriore del lobo temporale.

Zatorre e coll. (2004) svolgono uno studio PET per indagare i correlati neurali dell'elaborazione delle caratteristiche dell'oggetto uditivo. Gli autori utilizzano 45 suoni ambientali di diverse categorie semantiche (versi degli animali, elementi naturali, strumenti musicali, sirene, ecc) e i soggetti dovevano indicare da una scala da 1 a 10 quanto due suoni fossero simili (in cinque condizioni differenti). I risultati mostrano un'attivazione massima del lato superiore del solco temporale superiore anteriore (STS) destro, quando gli stimoli sono diversi, attivazione minima quando sono identici. Durante i compiti si è vista anche l'attivazione dell'area frontale inferiore destra. Gli autori concludono affermando che questi dati confermano l'importanza della via anteroventrale da un punto di vista funzionale per l'elaborazione dell'oggetto.

Lewis e coll. (2004) indagano i correlati neurali del riconoscimento dei suoni riconoscibili (versi animali, oggetti che cadono, strumenti, liquidi, ecc) e irriconoscibili tramite fMRI. Gli autori riscontrano l'importanza del ruolo semantico nell'attivazione di aree diverse. In particolare, gli autori riscontrano che i suoni riconoscibili attivano aree associate con l'elaborazione semantica, collocate principalmente nell'emisfero sinistro, ed anche una forte attività bilaterale nella parte posteriore del giro temporale mediano (pMTG). Gli autori aggiungono che l'attivazione bilaterale del pMTG coincide parzialmente con le aree implicate nell'elaborazione visiva di alto livello e riconoscimento di oggetti. Gli autori interpretano questi dati come prova che il pMTG è un'area di elaborazione multimodale sovra modale sulle informazioni circa gli oggetti e il movimento associato agli oggetti, il che potrebbe spiegare la conoscenza richiesta dal riconoscimento dei suoni (e la loro localizzazione) in ambienti familiari e l'agnosia uditiva per suoni ambientali riportata in alcuni studi su lesioni umane.

Bud e Timora (2013) riportano che nella modalità uditiva l'ampiezza modulata dai suoni elicita risposte SSR massime a 40 Hz (Galambos et al., 1981; Rees et al., 1986).

Studi successivi hanno mostrato anche un'attivazione di aree all'infuori della corteccia uditiva, tra cui la regione corticale prefrontale. Inoltre, recenti studi hanno mostrato che compiti di identificazione uditiva possono attivare aree considerate dorsali, e viceversa compiti di localizzazione spaziale hanno mostrato attivare aree ventrali. Ne consegue che un modello di elaborazione in parallelo rischia di essere riduttivo. Bizley e Cohen (2013) riportano alcuni esempi in merito alla complessità dell'elaborazione dell'oggetto uditivo: per esempio, quando viene presentata una sequenza ritmica con suono uguale da uno stesso punto, a seconda della

temporalità viene percepito come evento unico. In altro modo, compiti spaziali possono richiedere l'identificazione sonora anche se il compito è spaziale, e viceversa.

1.1 Premesse sul *binding* e *binding* audio-tattile.

Binding: definizione e aspetti generali.

In qualsiasi situazione quotidiana, in modo quasi continuo, i nostri sensi sono bombardati da diversi segnali sensoriali. Per avere una stima il più attendibile possibile delle proprietà degli oggetti e dell'ambiente che ci circonda, o almeno, quella meno variabile, abbiamo bisogno di combinare, integrare e definire in modo coerente il rumore che proviene dai diversi sensi, per definire cosa appartiene a uno stesso oggetto o a diversi oggetti (Spence, 2011).

Il termine *binding* si riferisce al processo attraverso cui diversi attributi di oggetti nell'ambiente (rappresentati in una sorta di matrice sensoriale) vengono unificati, integrati o mantenuti in modo distinto, in modo tale da avere un'interpretazione coerente del mondo attorno a noi. Senza questi meccanismi e processi legati all' integrazione percettiva che ci consentono di 'interpretare' gli input sensoriali, saremmo circondati da stimoli caotici o, in altre parole, dal rumore. Il problema (o i problemi) del *binding* si propone sia se si considerano modalità sensoriali differenti, sia se si considera una singola modalità sensoriale. Anche all'interno di una stessa modalità sensoriale infatti, vi sono delle caratteristiche che vengono codificate separatamente per poi essere integrate.

In altre parole, il *binding* è ciò che ci consente di fare una stima veritiera, o meno variabile del mondo che ci circonda e riguarda in modo pervasivo tutti gli aspetti della percezione della realtà che ci circonda.

Il problema o i problemi del *binding*: classificazioni.

L'espressione "*binding problem*" è stata introdotta per indicare le difficoltà che si possono verificare nel selezionare quelle risposte che vengono evocate da un singolo oggetto percettivo ed è un termine che si colloca tra la filosofia e le neuroscienze, assumendo nel primo caso il significato di "unità percettiva cosciente" e nel secondo caso di "percetto unitario". Ad oggi, il problema del binding non è stato ancora risolto, ed anzi, di recente è stato suggerito di riferirsi

al *binding problem* non più come ad un singolo problema, ma ad una classe di problemi, dal momento che è implicato in diversi fenomeni percettivi (Roskies 1999, Singer 1999, Scholz 2001, Holcombe 2009).

Una prima teorizzazione del '*binding problem*' è stata formulata da C. Von der Malsburg nel 1981, in ambito visivo. Le origini di questo problema, infatti, sono state analizzate in primis nell'ambito della modalità sensoriale visiva, probabilmente perché originariamente ritenuta come il canale sensoriale attraverso cui l'uomo trae la maggior parte delle informazioni sul mondo circostante (il 90% degli input percettivi è di tipo visivo; Bujacz and Strumillo, 2007), nonché la prima modalità ad essere largamente oggetto di studio.

Il primo esempio di *binding* è stato illustrato da Frank Rosenblatt nel 1961 con la formulazione di un ipotetico sistema percettivo basato su un *network* neurale classico. Questo sistema è costituito da quattro neuroni, due dei quali rispondono in presenza degli oggetti, uno risponde in presenza di un triangolo, l'altro in presenza di un quadrato; entrambi generalizzano sulla posizioni degli altri due neuroni che indicano la posizione degli oggetti (uno risponde alla parte superiore, l'altro alla parte inferiore). La generalizzazione di entrambi

dà la natura dell'oggetto. Se viene presentato un solo oggetto la risposta è adeguata, se invece vengono presentati simultaneamente due oggetti, due neuroni si attivano simultaneamente (per esempio quello che rileva il triangolo e quello che rileva la parte alta degli oggetti). Rosenblatt (1961) riporta quindi un esempio in cui l'assenza di una caratteristica da sola non è sufficiente per la rappresentazione di due oggetti simultanei (Stoet e Hommel, 2002).

Roskies (1999) specifica che il motivo per cui ci si riferisce al *binding* come un problema è deducibile da due indicatori principali che indicano che il cervello presenta una difficoltà nell'integrazione di caratteristiche: 1) un primo indicatore sono le congiunzioni illusorie, 2) ed un secondo indicatore sono le mancate integrazioni presenti in alcuni danni cerebrali. Egli è uno dei primi ad evidenziare come questo termine di fatto possa includere diversi fenomeni e che possa riguardare altre modalità sensoriali, oltre alla visione. Egli propone una distinzione in **problemi di *binding* percettivo**, che riguardano gli aspetti unificanti di un percetto (per esempio, indica che possa esistere un *binding* uditivo che consente di discriminare una voce in mezzo alla folla, un *binding* temporale che possa consentirci di rilevare il movimento degli oggetti, un *binding* crossmodale che riguarda l'integrazione di diverse modalità come diversi aspetti di un unico evento) e **problemi di *binding* cognitivo**, che sono relativi alla rappresentazione concettuale di un percetto, alla identificazione crossmodale e alla ricostruzione mnestica di un oggetto (per esempio, la rappresentazione di una mela viene collegata ad una serie di conoscenze semantiche relative a diverse sue caratteristiche come il

sapore e il colore, o riconoscere un oggetto e codificarlo un una rappresentazione strutturata e unificata). L'autore specifica che questa distinzione è puramente fittizia, ma serve a rendere l'idea sul fatto che il binding si ripropone per diversi tipi di processi cerebrali.

Scholz (2001) riprende la distinzione proposta da Roskies ampliandola ulteriormente distinguendo ogni singola modalità sensoriale, diverse modalità sensoriali e aggiungendo l'integrazione senso-motoria (relativa alla rappresentazione sensoriale di un oggetto e il relativo comando motorio per agire in relazione a quell'oggetto). Egli nel binding percettivo include il *binding* visivo, uditivo, nel tempo e cross-modale e nel *binding* cognitivo include quello che correla un concetto ad un percetto, l'identificazione cross-modale, aggiunge il ***binding* senso-motorio** e la ricostruzione della memoria.

Hommel (2004) e Stoet e Hommel (1999, 2002) descrivono il problema del *binding* all'interno del controllo dell'azione. Similmente a quanto accade nella percezione, la rappresentazione simultanea di azioni multiple richiede un meccanismo per la codifica in cui le caratteristiche motorie vengono assemblate insieme, ovvero è presente una molteplicità relativa alle componenti dell'azione, le caratteristiche dell'azione e i sistemi di controllo dell'azione. Hommel (1998) e Hommel e Milliken (2007) riportano il problema del *binding* tra percezione e azione, denominato **binding stimolo-risposta**, ipotizzando che ciò che percepiamo e le azioni che svolgiamo devono essere necessariamente collegate l'un l'altra.

Zimmer e coll (2006) indicano con l'espressione di ***binding* di memoria** l'integrazione di un qualsiasi caratteristiche possibili e contesti (per esempio colore, suono, emozione, attivazione, contesto) e concerne il modo in cui ricordiamo le associazioni tra i diversi elementi di un evento e come il cervello crea e mantiene queste associazioni.

Holcombe (2009) distingue due tipi di *binding*: ***binding* spaziale** e ***binding* temporale**. In particolare, egli afferma che l'esistenza del *binding* spaziale viene confermata dai compiti utilizzati per la *Feature Integration Theory* della Treisman (1996) di ricerca visiva. Infatti, nel momento in cui i soggetti devono identificare un target che richiede un'integrazione di caratteristiche impiegano più tempo, ciò si verifica perché il *binding* richiede attenzione e l'attenzione si ipotizza che si muova da un posto all'altro alla volta, integrando le caratteristiche di volta in volta finchè il target non viene trovato. Un altro esempio indicato dall'autore, sono le congiunzioni illusorie (Treisman e Schmidt, 1982) che vengono interpretate come conseguenti a un fallimento del *binding*. Holcombe (2009) però, aggiunge che anche nei casi in cui l'attenzione è presente, il *binding* può fallire lo stesso, per esempio, in alcune sindromi neuropsicologiche in cui le difficoltà di localizzazione corroborano l'ipotesi che il binding richieda l'elaborazione della localizzazione (*location tags*). Per quanto riguarda il *binding*

temporale, l'autore riporta che l'elaborazione del *binding* è più lenta relativamente all'identificazione di caratteristiche. L'autore riporta diversi esempi di *binding* visivo dei tempi di integrazione, e conclude, in base alla risoluzione temporale, che il *binding* richiede più tempo rispetto all'identificazione di singole caratteristiche. L'autore aggiunge che bisogna considerare che diverse caratteristiche hanno diversa latenza e tempi di elaborazione.

Schmidt (2009) definisce i termini del problema del *binding* relativamente a: 1) come il sistema determina quali caratteristiche vanno integrate insieme e dovrebbero appartenere ad un singolo oggetto; 2) Come ciò viene segnalato nel cervello, nonché come la rappresentazione di caratteristiche dello stimolo integrate si differenzia rispetto a quelle non integrate; 3) come questa informazione viene riconosciuta dal resto del sistema, ovvero come il cervello riconosce e recupera una rappresentazione dello stimolo integrata una volta che le caratteristiche sono state integrate. Secondo l'autore una teoria che offra una soluzione al problema del *binding* dovrebbe rispondere a queste domande, alla costanza e raggruppamento degli oggetti nella stessa posizione. Egli indica diversi tipi di *binding*: *binding* spaziale e temporale (già descritti), problemi del *binding* presenti nella sinestesia e nella sindrome di Balint-Holmes, nell'integrazione multisensoriale, nel controllo dell'azione e il suo monitoraggio. Inoltre, riporta alcune delle più importanti teorie che hanno cercato di risolvere alcuni dei problemi ad esso collegati ed ogni teoria riporta un diverso aspetto del problema (o dei problemi) del *binding*: *binding* dovuto alla convergenza, *binding* dovuto alla codifica della popolazione, *binding* dovuto alla sincronizzazione, *binding* dovuto all'attenzione, *binding* dovuto all'elaborazione ripetuta, *binding* e consapevolezza.

Di Lollo (2012) risponde alla definizione del *binding* data Wolfe (2012), che collega il problema del *binding* al limite dell'attenzione selettiva (il problema del *binding* è il problema di conoscere come le caratteristiche base sono collegate l'un l'altra). Egli considera che il problema del *binding* di caratteristiche così come descritto da Von der Malsburg è formulato male. Aggiunge dicendo che il termine '*binding*' indica molte cose. Riprende la *Feature integration Theory* (Treisman, 1996) che identifica non meno di 7 diversi binding problems (dal *property binding* al *conditional binding*) e il fatto che più di recente il *binding* è stato applicato a problemi riguardanti la percezione, l'attenzione, la memoria di lavoro e la memoria a lungo termine (Brockmole e Franconeri, 2009). In questo senso, per l'autore risolvere il problema del *binding* è futile perché significherebbe porre soluzione a diversi problemi e questa soluzione potrebbe non essere applicabile per gli altri. Aggiunge che bisogna porre una distinzione chiara tra 'problema nella percezione dell'oggetto' e problema di 'integrazione delle caratteristiche' (feature binding).

Feldman (2013) riporta il "**Neural Binding Problem (NBP)**" come comprendente almeno quattro problemi distinti , con diversi pre-requisiti computazionali e neurali: considerazioni generali sulla coordinazione, l'unità soggettiva della percezione, *binding* di caratteristiche visive, *binding* variabile. L'autore afferma che il 'problema del *binding*' riguarda la codifica di diversi aspetti degli oggetti legati a diversi circuiti cerebrali che sono codificati da diversi circuiti neurali che possono essere combinati tra loro a seconda che si tratti di: percezione, decisione o azione (Brockmole e Franconeri, 2009). Secondo l'autore ci sono stati dei progressi nella comprensione dei substrati neurali per la coordinazione nel cervello, ma c'è ancora un velo di mistero dovuto al fatto che si tratta di diverse questioni relative allo stesso 'problema'.
1) Le 'considerazioni generali sulla coordinazione' indicano gli studi relativi alla sincronizzazione temporale e riguardanti come il cervello coordina circuiti multipli e come si sviluppano. Gli oggetti e attività che si verificano nello stesso tempo e spazio sono visti come unitari. 2) L'unità soggettiva della percezione' è relativa al fenomeno della percezione unitaria. Un esempio è il problema mente-corpo che si ricollega agli studi sui correlati neurali della consapevolezza (*consciousness*).
3) Il '*binding* di caratteristiche visive' riprende il paradigma di ricerca visiva della Treisman (1999) (come riusciamo a distinguere un quadrato rosso dal cerchio blu rispetto ad un quadrato blu e un cerchio rosso).
4) Il 'binding variabile' riguarda come il cervello impiega ogni fetta temporale per ciascuna variabile ('*temporal phase synchrony'*), nonché come due caratteristiche vengono congiunte sebbene non siano note da prima. L'autore ricollega questo problema con l'analisi linguistica e il pensiero simbolico (per esempio le parole di una frase devono essere integrate insieme per dar valore ad una frase ed essere capite o per capire cosa intende per Unificazione di caratteristiche, l'aggettivo deve concordare per genere e numero con il sostantivo).
Altri contesti, differenti rispetto a quelli elencati finora, riguardano il *binding* relativo alla comprensione linguistica (Jackendorff, 2002 e van der Velde e de Kamps, 2006) e *binding* relativo al ragionamento (Shastri e Ajjanagadde, 1993). Solms e Turnbull (2002), Singer (2001) e Mashour (2004) si riferiscono al *binding* collegandolo all'esperienza cosciente (*consciousness*).

Origine e soluzione del *binding*.

Un'ipotesi circa l'origine del *binding*, ovvero del perché sia necessario questo fenomeno, è quella accennata nell'introduzione al problema: il *binding* costituisce una sorta di soluzione dei conflitti tra modalità sensoriali o a stimoli ambigui, altrimenti saremmo circondati dal rumore.

Tra i primi studi in quest'ottica, ricordiamo gli studi sulla **multistabilità**, ovvero le ricerche sulla percezione instabile nelle differenti modalità sensoriali. Schwartz e coll. (2012) fanno riferimento alla multistabilità quando un singolo stimolo fisico produce alternanze tra diversi percetti soggettivi, o quando gli stimoli hanno più di una organizzazione percettiva plausibile e quando queste organizzazioni non sono compatibili l'una con l'altra. Secondo gli autori, nella maggior parte dei casi, la multistabilità si basa sulla competizione nella selezione e nel *binding* dell'informazione dello stimolo. In questa accezione, il *binding* fa riferimento al processo per mezzo del quale i diversi attributi di un oggetto nell'ambiente (come rappresentato in una matrice sensoriale) sono collegati insieme entro i nostri sistemi percettivi, in moda da fornire una coerente interpretazione del mondo circostante. Secondo gli autori, la multi stabilità può essere usata come metodo per studiare i processi di binding entro e attraverso le modalità sensoriali.

Ogni oggetto (come già detto, storicamente è stato analizzato per primo l'oggetto visivo) presenta una propria forma associata ad una specifica localizzazione, e queste due caratteristiche di un oggetto vengono integrate in modo tale che noi possiamo ottenere una rappresentazione unificata di quell'oggetto. É proprio questa implicita associazione di caratteristiche che ci consente di identificare un oggetto come distinto e ci consente, per esempio, di non incappare nelle congiunzioni illusorie.

Nato come problema teorico, è stato poi studiato anche tramite diversi studi psicofisiologici e di neuroimmagine, in cui vengono riportati risultati che enfatizzano una sorta di segregazione per l'elaborazione di caratteristiche diverse (soprattutto in ambito anatomico) ma anche studi che enfatizzano una comunicazione tra aree cerebrali diverse. Studi di neuroimmagine sui correlati neurali annessi (tecniche: ERPs, EEG, tDCS, TMS, MEG, fMRI): mostrano diverse ipotesi a livello neuronale che possono riguardare meccanismi di sovrapposizione, iperconnessione e disinibizione dei feedback.

Gray (1999) ritiene che il *binding* non sia un problema per il sistema nervoso, dal momento che l'evoluzione ha modellato la sua organizzazione per risolvere il problema nel modo più efficiente possibile. Secondo l'autore si tratta invece di un problema per coloro che provano a capire come il sistema nervoso realizza un compito. Diversi studi riportano il problema del *binding* in cervelli sani. Per esempio, le già citate congiunzioni illusorie riportate da Treisman e Gelade (1980) e Wolfe e Cave (1999) vengono riportate come evidenza di un problema di *binding* nel cervello. Altri esempi riguardano fenomeni sinestetici come riportato da Robertson (2003). Come già accennato, problemi del *binding* si verificano anche in presenza di danni

cerebrali, per lo più coinvolgenti il lobo parietale, come per la sindrome di Balint, ovvero l'incapacità di percepire più di un oggetto alla volta.

Come per la sua formulazione, anche le soluzioni legate al problema del *binding* sono state formulate e verificate inizialmente studiando la modalità visiva.
La soluzione più semplice è stata proposta negli anni '50 Horace Barlow (per es. 1960) propone la **'dottrina del neurone'**, per cui ogni 'cellula cardinale' serve come detettore di ogni combinazione di oggetti, ovvero l'ipotesi che vi sia un singolo neurone (o unità rappresentazionale) per ogni possibile combinazione di caratteristiche. Successivamente, David Hubel e Torsten Wiesel (1962) propongono che cellule con complessa selettività allo stimolo possono essere formate dalla convergenza di cellule con proprietà semplici ed ottengono evidenze in merito. Alcune di queste cellule, sensibili a rispondere solo in presenza di forme complesse o volti, vennero chiamate in seguito 'cellule nonne' o 'cellule gialle Volkswagen' (visto che rispondono quando si vede la nonna o la sua vecchia macchina gialla). Dal momento che potenzialmente esiste una illimitata possibilità di combinazioni, e dal momento che dovrebbero esserci dei sistemi di riadattamento e di ricodifica per classificare oggetti nuovi, questa ipotesi è stata criticata e abbandonata, sebbene in alcuni casi è stata verificata la presenza di sistemi neuronali selettivi per una certa combinazione di caratteristiche, capaci di risolvere il problema del *binding* in alcuni casi (Risenhuber e Poggio, 1999). Questa teoria è stata ribattezzata da Schmidt (2009) come teoria del ***binding* per convergenza**.
Un'evoluzione successiva all'ipotesi del binding per convergenza, è stata data ipotizzando l'esistenza di una rete di cellule in cui ogni cellula può essere ON o OFF. Secondo questa teoria, è possibile assegnare un oggetto per ogni possibile stato dell'intera rete. Questo schema è stato denominato 'population coding', da cui il nome assegnato da Smidth (2009) di teoria del ***binding* per codifica della popolazione**. Ad ogni modo, questa teoria non supera il problema della teoria del *binding* per convergenza, dal momento in cui si sostituisce la codifica della cellula cardinale con la popolazione, la teoria è in grado di spiegare come il *binding* è segnalato, ma non come viene stabilito, ovvero come queste proprietà vengono stabilite e generate inizialmente e come viene riconosciuta dal sistema. Ragion per cui queste due teorie spiegano più come avviene la rappresentazione di un oggetto piuttosto che risolvere il problema del *binding*, sebbene possano spiegare in parte il *binding* spazio-temporale, visto che le cellule codificano diverse parti dell'oggetto e loro posizione (Smidth, 2009).

Negli anni '80 Christoph von der Malsburg suggerì che cellule che codificano le proprietà di uno stesso oggetto possono scaricare in sincronia (egli si riferiva a visivo, riportando una finestra temporale di 5-10 ms). Wolf Singer (1999) è il maggior esponente di questa teoria. Egli ipotizza la presenza di una sincronizzazione ritmica di attivazione neuronale selettiva per caratteristiche appaiate. Sebbene l'ipotesi sia stata corroborata, la relazione tra questo fenomeno e il problema del binding rimane ancora da chiarire. Esistono infatti alcuni aspetti controversi, come per esempio come avviene la sincronizzazione top-down di queste cellule, e più in generale come vengono stabiliti i meccanismi di scarica simultanea (più che la scarica sincrona per sé). In qualche modo questa teoria amplia il binding per codifica della popolazione, incappando nelle stesse lacune esplicative (Smidht, 2009). Tutte le teorie esposte fino adesso spiegano come il *binding* possa essere segnalato, ma non come viene generato e decodificato dal sistema, nonché nella spiegazione del fenomeno stesso.

Tornando indietro storicamente, la proposta più conosciuta per la soluzione del *Binding problem*, nonché una delle teorie più influenti per l'attenzione, e la più volte accennata, è la "*Feature Integration Theory*" (Treisman e Gelade, 1980; Treisman e Schmidt, 1982), la quale prevede che il binding dipenda dall'attenzione selettiva e si sviluppi secondo uno stadio di elaborazione preattentiva (in parallelo) in moduli indipendenti e un secondo stadio attentivo di integrazione mediata dall’attenzione (sequenziale). Treisman assume che il sistema visivo sia equipaggiato da 'mappe di caratteristiche', e da mappe diverse per ogni caratteristica possibile. Queste 'mappe di caratteristiche' sono connesse a una 'mappa esperta di localizzazione' che codifica la posizione degli oggetti nel campo visivo. Il *binding* si verifica nel momento in cui l'attenzione viene indirizzata ed è quindi strettamente legato all'attenzione selettiva. Questa teoria è stata recentemente ampliata e ad oggi esistono altri modelli in cui il la prestazione in compiti di ricerca visiva non viene strettamente legata al ruolo dell'attenzione e viene fatta dipendere da altri fattori (es. in ambito visivo: meccanismi di segmentazione dell'immagine, l'elaborazione della salienza locale, masking, ecc). La stessa distinzione in elaborazione parallela vs sequenziale è stata messa in discussione. Ad ogni modo, questa teoria non spiega il problema del *binding* quando le diverse caratteristiche si trovano nella stessa posizione.

Lamme e Roelfsema (2000) asseriscono che per comprendere il binding bisogna fare riferimento alle connessioni di feedback deputate all'attività ripetuta tra le diverse aree collegate gerarchicamente nell'elaborazione dell'informazione in entrata. Tramite l'analisi della latenza di risposta delle diverse aree corticali (utilizzano stimoli visivi), gli autori affermano che ogni stimolo crea un'onda di attivazione che viaggia dalle aree posteriori a quelle anteriori. Dal momento che tra un'attivazione e l'altra c'è un tempo di latenza, gli autori deducono che vi

sia poco o quasi nessun periodo di tempo per integrare il feedback dalle altre cellule e ipotizzano che la prima onda dell'attivazione visiva viaggia attraverso il sistema come una '*feed-forward*' la quale è essenzialmente libera dall'informazione di feedback intracorticale. Gli unici processi di *binding* possibili avvengono nella fase di attivazione *feed-forward* per convergenza o per codifica della popolazione. Gli autori affermano che l'elaborazione *feedforward* costituisce solo un pre-requisito per il *binding* e che non è esaustiva nel risolvere il problema. Di contro, affermano che l'elaborazione ricorrente potrebbe spiegare il problema nei termini in cui i detettori di caratteristiche (o popolazioni) attivati dal ' *feedforward*' inviano indietro le proiezioni rientranti, cioè alle aree precoci, cosicché l'attivazione di ciascun area può essere immediatamente ricollegata alle mappe topografiche precoci. In questo modo, le caratteristiche di un singolo oggetto possono essere elaborate in aree separate, ma allo stesso tempo possono riferirsi indietro alla singola localizzazione spaziale dalle interazioni ricorrenti con una mappa spaziale comune. L'analogia con la FIT della Treisman è evidente e gli autori fanno riferimento all'attenzione visiva selettiva come una forma specifica di attività ricorrente.

Aspetti temporali del *binding*.

Come già accennato, il primo a definire il termine di binding è stato teorizzato da von der Malsgurd del 1981 nell'ambito della modalità visiva. Abbandonando la teoria delle cosiddette "cellule cardinali" o "cellule nonne" secondo la quale ogni cellula è deputata all'elaborazione di una determinata combinazione di caratteristiche, per altro teoria molto improbabile dal momento che il numero delle combinazioni possibili (tendente a infinito) eccede la disponibilità del numero dei neuroni esistenti, von der Malsgurg propose l'esistenza di un duplice segnale: a) la stimolazione effettiva, trasmessa tramite un codice convenzionale, e b) l'attivazione sincronizzata di un gruppo di neuroni. L'attività dei neuroni di basso livello viene combinata solo quando il loro picco di attività è sincronizzato in pochi millisecondi per creare un assemblamento cellulare attivo sincronizzato (Shadlen e Movshon, 1999)

La sincronizzazione di queste cellule è dinamicamente modulata. Ciò significa che ogni singola cellula può appartenere a più di un assembramento neurale, risolvendo così il problema delle combinazioni arbitrarie e il numero limitato di neuroni. Egli quindi propose l'ipotesi di un *binding* temporale, domandandosi come il *binding* si possa realizzare.

Questo modello è stato applicato successivamente da Eckhorn et al (1988) e da Signer e coll. (1993, 1994, 1999a) ed ulteriormente estesa proponendo che la popolazione di cellule di basso livello sincronizza la loro attività in risposta a diversi elementi.

Una delle critiche che viene fatta a questo modello di *binding* temporale, proposta da von der Malsburg ed ampliata da Singer, è che pensare ad una sincronizzazione di gruppi neuronali e il non spiegare quale neurone e perché viene attivato in un gruppo neurale o un altro, riporta fondamentalmente il problema delle cellule cardinali ad un più ampio raggio (Shadlen e Movshon, 1999).

Successivamente è stata riscontrata anche l'importanza dell'analisi di alto livello, per esempio il ruolo della memoria nel ricostruire la struttura di oggetti o di scene già esperite. (Shalden e Movshon, 1999). Anche la letteratura neurologica supporta l'idea che il *binding* sia un processo ad alto livello (es. la sindrome di Balint presenta delle grossi deficit di *binding* visivo).

Se si riporta il problema del *binding* in termini operazionali, la domanda cambia, per cui non è più importante se i neuroni di basso livello ricevono input da poche fonti selezionate, ma quale di questi input veicola segnali importanti per una secondo stadio di elaborazione. Ne deriva che l'ipotesi del *binding* temporale non riesce a spiegare come la sincronizzazione possa essere utile per capire quale sia il criterio determinante per stabilire cosa viene precluso o meno della complessità del mondo circostante.

Ne segue che determinati picchi di attività devono veicolare una sorta di informazione addizionale che richiede il *binding* ma porsi questa domanda implica il domandarsi sul numero degli input, sul grado di *spikes* e sulla finestra temporale in cui si suppone la sincronizzazione veicoli un significato particolare. Inoltre, dovrebbero essere presi in considerazione anche altri fattori che influenzano l'efficacia dell'input sinaptico, come l'adattamento, l'amplificazione dendritica, ecc. Si consideri inoltre, che l'attività sincronizzata è per definizione oscillatoria, dal momento che una volta raggiunto un picco, ciascun neurone ha un determinato periodo di latenza.

L'importanza della risposta oscillatoria interneuronale è stata ben documentata da Köning e coll. nel 1995. Gli autori dimostrano che mentre correlazioni tr neuroni vicini possono essere riscontrati quando i neuroni emettono delle risposte senza motivo, correlazioni tra neuroni di regioni molto distanti sembrano essere raramente osservati in assenza di risposte oscillatorie.

Un altro tentativo di spiegazione del *binding* riguarda l'ipotesi del '*binding* senza sincronizzazione' (Olshausen et al., 1993, 1995; Tsotsos, 1995) la quale postula la convergenza delle cellule nelle aree corticali di alto livello con la consapevolezza che alcuni problemi non possono essre risolti da una semplice convergenza. Sempre formulate in ambito visivo, e con riferimento principalmente al binding di caratteristiche, queste teorie considerano il problema del *binding* come implicitamente risolto nelle aree corticali superiori che ricevono e combinano in modo dinamico l'informazione.

Shadlen e Movshon (1999) propongono una soluzione al *binding* nel 'contesto dell'azione', ovvero il *binding* è specificamente correlato ai circuiti neurali che programmano il comportamento, ovvero come raggiungere un oggetto. Il *binding*, quindi, è visto come un raggruppamento di caratteristiche che guidano al bersaglio comportamentale, con particolare riferimento all'importanza della corteccia parietale nel *binding*.
Engel e Singer (2001) riportano studi animali e umani discutendo sull'importanza dei meccanismi di *binding* temporale per la consapevolezza sensoriale. Gli autori argomentano l'importanza della sincronizzazione neurale nel giro e come essa nel giro di secondi sia cruciale per l'elaborazione sensoriale consapevole, soprattutto con l'oscillazione veloce delle frequenze del raggio gamma. Ipotizzano inoltre che sia coinvolta anche nell'attivazione o *arousal*, nell'integrazione percettiva, nell'attenzione selettiva e nella memoria di lavoro.

Come già accennato, i nostri sensi sono bombardati da segnali uni sensoriali differenti, in ogni momento e per ottenere la più veridica e meno variabile stima delle proprietà ambientali. Così combiniamo le informazioni unisensoriali (ignorandone il rumore) e le riferiamo allo stesso oggetto, o consideriamo come appartenenti a diversi oggetti o eventi separati (Spence, 2011). Inoltre, l'integrazione multisensoriale spesso comporta un vantaggio in termini di prestazione e percezione che riflette l'informazione aggiuntiva data dall'informazione ridondante o complementare dei vari sensi (Wallace e Stevenson, 2014). Il nostro cervello continuamente integra l'informazione proveniente dai vari canali sensoriali per creare dei percetti solidi circa gli oggetti e gli eventi. Come affermano Stein e Meredith (1993) questi pezzi di informazione sono in qualche modo complementare e si completano l'un l'altro, incrementando in tal modo l'affidabilità della nostra percezione.

Binding multisensoriale o crossmodale: aspetti generali

Come già detto, per lungo tempo le diverse modalità sensoriali sono state indagate separatamente e considerate indipendenti, in quanto deputate all'analisi di forme di energia fisica specifiche (quali la luce per la vista o la pressione cutanea per il tatto o le onde sonore per l'udito e così via) (Galloni, 2003). Perfino la caratteristica modulare che Fodor (1983) definisce 'incapsulamento informazionale' era stata teorizzata sulla base dell'assunto dell'isolamento delle modalità percettive, che egli chiamava trasduttori, il cui solo *output* veniva ad interagire con gli *output* delle altre modalità ai fini della generazione di una rappresentazione o dello svolgimento di un compito cognitivo di alto livello (Fodor, 1983). In quest'ottica, anche nel modello additivo semplice di Anderson del 1974 una giunzione di

risposta bimodale viene considerata come la media pesata delle condizioni unimodali (viene fatta una stima del peso relativo di ogni input per ciascuna modalità nel percetto congiunto).
Sebbene, quindi, inizialmente si sia preferito studiare gli oggetti considerando solo le proprietà unisensoriali, trattandoli come se fossero degli oggetti uni sensoriali di per sé, in realtà ciò che percepiamo costantemente e che ci circonda sono oggetti multisensoriali, o unità percettive complesse in quanto costituite da informazioni provenienti da modalità sensoriali diverse.
Da un punto di vista storico, è importante rilevare come già Müller (1838), nel teorizzare la cosiddetta 'legge dell'energia nervosa specifica', avesse citato l'effetto ventriloquo (un caso puro di integrazione cross-modale) come un'illusione percettiva che presenterebbe un'eccezione rispetto alla sua legge. Nonostante il fatto che, a partire da allora, vi siano stati sporadici lavori sulle 'illusioni multisensoriali', si dovrà attendere la seconda metà del Novecento per arrivare ad un movimento realmente sistematico di ricerche sulla percezione cross-modale.
In sintesi, quindi, il problema del *binding* multisensoriale può essere posto come domanda circa le condizioni in cui il sistema percettivo unisce due o più stimoli provenienti da modalità sensoriali diverse (percetti parziali), a far parte di un solo oggetto, di una sola unità complessa (percetto finale unitario) e qual è l'aspetto di questa unità multisensoriale rispetto alle sue componenti unisensoriali.

Studi comportamentali e psicofisici

Gli studi sul problema del *binding* multisensoriale, o crossmodale, includono una serie di problemi, tra cui i ruoli che fattori temporali e spaziali giocano nel modulare l'integrazione multisensoriale e la domanda circa le corrispondenze cross modali tra le diverse caratteristiche uni sensoriali (es. altezza uditiva e la grandezza visiva), ovvero corrispondenza tra caratteristiche degli stimoli in diverse modalità sensoriali. La letteratura suggerisce che le corrispondenze cross modali devono essere considerate lungo una congruenza semantica e spaziotemporale. (Spence, 2011)
Il problema del *binding* multisensoriale è stato studiato inizialmente a partire da alcune considerazioni tratte da studi sperimentali sulle reazioni del sistema percettivo posto di fronte a stimoli cosiddetti conflittuali, o meglio: discordanti. La particolarità di questo tipo di studi è quella di permetterci di analizzare il funzionamento del sistema percettivo senza la presunzione che esso stia valutando un solo oggetto costituito da informazioni conflittuali (Pasquinelli, 2003). Pasquinelli (2003) afferma che questi studi ci mettono di fronte a quella che appare come una presa di decisione del sistema percettivo quanto alla formazione o meno di unità

multisensoriali più complesse e quanto all'aspetto di queste ultime. Invece che un semplice ricevitore passivo di stimoli il sistema percettivo ci appare dunque come un agente costruttivo che opera scelte a partire da blocchi da costruzione più o meno elaborati.

Talsma e coll (2010) studiano le interazioni tra l'attenzione e l'integrazione multisensoriale. Essi affermano che l'integrazione multisensoriale avviene in diversi stadi dell'elaborazione dell'info che sono collegate e possono essere modulate dall'attenzione. Meccanismi *stimulus-driven* e *bottom-up* indotte dalle interazioni cross modali possono automaticamente catturare l'attenzione attraverso eventi multisensoriali, specialmente quando la competizione verso il focus in altri posti è relativamente bassa. Inversamente, l'attenzione top-down può facilitare l'integrazione degli input multisensoriali e guidare alla diffusione dell'attenzione attraverso le modalità sensoriali. Sembra nell'ambito dell'integrazione multisensoriale e attenzione cross modale, Koelewijn e coll (2010) ipotizzano l'esistenza di un frame work parallelo di elaborazione in cui entrambi possono prendere parte e interagire a livelli multipli, non completamente indipendenti e non automatici.

Shams e coll. (2011) studiano come l'esperienza multisensoriale influenza l'elaborazione uni sensoriale successiva. Essi ipotizzano che la modificazione continua delle rappresentazioni uni sensoriali tramite le relazioni multisensoriali potrebbe essere una strategia generale di apprendimento utilizzata dal cervello.

Wallace e Stevenson (2014) affermano che la relazione spaziale e temporale dell'informazione multisensoriale provvede un'informazione su come i vari sensi possono essere 'legati' ('bound') per creare una rappresentazione unitaria. In particolare, gli autori si focalizzano sugli aspetti temporali dell'integrazione multisensoriale e sulla cosiddetta 'finestra temporale del binding multisensoriale' (TWB, '*multisensory temporal binding window'*) , sebbene ribadiscono che entrambi i fattori, spaziale e spazio-temporale giocano nella costruzione multisensoriale di una Gestalt (intesa come 'buona forma') e come questi possano influire nell'interazione multisensoriale a livello neurale, comportamentale e percettivo, agendo come dei 'filtri' del sistema multisensoriale (es. Bell et al., 2001; Diederich et al., 2003; Neil et al., 2006; Royal et al., 2009). Come già ripetuto più volte, esperire il mondo in modo coerente implica l'integrazione dell'informazione proveniente dalle varie modalità sensoriali e la percezione unificata degli oggetti sia in termini spaziali che temporali **(**Zmigrod e Leor Zmigrod, 2015**)**. Per quanto riguarda l'aspetto temporale è stata riscontrata una cosiddetta finestra temporale entro cui un singolo evento percettivo percepito come distinto e indipendente (per esempio de Gelder e Bertelson, 2003). Ciò diventa particolarmente rilevante quando vi sono discrepanze tra gli attributi fisici, la trasduzione sensoriale e i meccanismi

neurali coinvolti nell'elaborazione di informazioni provenienti da diversi sensi (Vroomen e Keetels, 2010). Di conseguenza, il sistema percettivo è costruito per adattarsi ad un raggio limitato di asincronie temporali di input multisensoriali: da queste considerazioni è nata l'idea dell'esistenza di una finestra temporale che sostiene la percezione dell'integrazione multisensoriale (per esempio Spence e Squire, 2003). In altre parole, esiste un'alta probabilità che stimoli di modalità sensoriali siano integrati in un singolo percetto multisensoriale entro questa finestra temporale del binding multisensoriale, malgrado l'assoluta mancanza di simultaneità temporale. Studi sui fattori temporali che influenzano la percezione degli stimoli multisensoriali sono stati svolti da Dixon e Spittz nel 1980, con stimoli di tipo audiovisivo con un paradigma di giudizi temporali. I soggetti dovevano indicare quando secondo loro due stimoli (uno visivo e uno uditivo) fossero simultanei o asincroni (gli stimoli sono presentati variando l'asincronia dell'*onset* dello stimolo; SOAs). I soggetti riportavano un'alta percentuale di stimoli simultanei anche con un ritardo di centinaia di millisecondi. Questi risultati suggeriscono che sia presente un certo grado di tolleranza temporale nella percezione della simultaneità degli stimoli multisensoriali asincroni, creando una sorta di 'finestra' temporale in cui stimoli multisensoriali sono percepiti legati o integrati (es. Stevenson e Wallace, 2013). Secondo Wallace e Stevenson (2014) la finestra temporale del *binding* multisensoriale (TBW) è un meccanismo altamente adattivo, dal momento che permette che l'informazione multisensoriale sia legata insieme e anche nel momento in cui origina a diverse distanze dal soggetto. Nel caso dell'audiovisivo questa utilità biologica è basata sulle differenze sostanziali della propagazione temporale dell'energia visiva e uditiva. Un'altra evidenza dell'importanza del ritardo biologico viene vista nelle misure del punto di simultaneità soggettiva (PSS), l'*offset* temporale esatto (misurato all'organo sensoriale) in cui gli individui percepiscono gli *inputs* come sincroni. Studi recenti mostrano caratteristiche salienti sulla TBW riguardanti il fatto che la finestra differisce nella dimensione per stimoli differenti, in base ai diversi contesti e stimoli (es, in audiovisivo: Stevenson & Wallace, 2013; van Eijk et al., 2008; Vatakis & Spence, 2006). Inoltre, la TBW mostra un consistente grado di variabilità da soggetto in soggetto (es. in audiovisivo Stevenson, Zemtsov, & Wallace, 2012) e continua ad aumentare con lo sviluppo (es. Hillock et al., 2011). Infine, è stato dimostrato che la TBW è molto malleabile in modi diversi, sia aggiustando le statistiche temporali dell'ambiente (es in audiovisivo Fujisaki et al., 2004), e in studi sulla plasticità percettiva hanno mostrato che può essere sostanzialmente ristretta con *training* di *feedback* (Powers et al., 2009, 2012; Stevenson et al., 2013; Schlesinger et al., 2014).

Recentemente, si è ipotizzato un modo diverso di pensare questi fattori, ovvero verso una visione più dinamica e contestuale dell'applicabilità di questi principi (Doehrmann & Naumer, 2008; van Atteveldt et al., 2014). Per esempio, ognuno di questi fattori avrà 'peso' diverso a seconda dei diversi contesti e ci sarà un set specifico di flessibilità di ruoli interattivi o principi correlati alla performance che in qualche modo detterà il prodotto finale della combinazione degli stimoli multisensoriali.

Spence (2010) reputa che lo sviluppo di modelli computazionali può dire qualcosa o addirittura predire quale effetto avrà la presentazione di uno stimolo in ciascuna modalità sulla percezione della persona di uno stimolo presentato in un'altra modalità. Per esempio, il modello bayesiano assume crescente rilevanza nel risolvere il problema del *binding*. Anche Shams e Kim (2010) nella loro revisione, constatano che i ricercatori stanno iniziando a sviluppare modelli computazionali più avanzati che sono capaci di spiegare come il problema del *binding* possa verificarsi in situazioni in cui stimoli multipli sono presentati contemporaneamente (Wozny e coll, 2008).

Ad oggi, viene riconosciuta l'importanza di studiare l'oggetto considerando l'interazione tra le modalità sensoriali, considerando quindi l'oggetto multisensoriale come risultato dell'interazione tra le varie modalità sensoriali. Come affermano Alais e coll. (2010) non è possibile capire il mondo isolando i sensi, perché tutti i sensi interagiscono tra loro. Gli autori affermano che la corteccia cerebrale è fondamentalmente multisensoriale. I vantaggi di un sistema percettivo multisensoriale sono due, ovvero che i sensi possono identificare segnali da un singolo evento (es. luce o energia del suono) e che possono monitorare regioni complementari dello spazio (es. detezione uditiva dietro la testa dove non c'è nessuna rappresentazione visiva). Così come detto da Shams e Kim (2010) , la ridondanza può migliorare la percezione in molti modi. L'integrazione non consapevole dei segnali attraverso modalità porta a un accoppiamento inutile che ostacola la percezione, piuttosto che aiutarla. La sfida centrale in un sistema multisensoriale, dunque, è accettarsi quali segnali integrare e quali segregare: dilemma noto, appunto, come '*binding* problem'. Motivo per cui, in molti casi gli studi sulla ricerca multisensoriale presentano segnali semplici isolati. La combinazione dei segnali e le leggi di corrispondenza spaziale e temporale hanno fornito una via di risposta al *binding*.

Integrazione a livello neuronale.

Originariamente si presumeva che l'integrazione tra le informazioni elaborate indipendentemente dalle diverse modalità avvenisse in aree corticali separate, che le regioni sottocorticali fossero specializzate nell'integrazione delle distinte modalità che che questo tipo di computazioni venissero elaborate gerarchicamente tardivamente (Fellemann e Van Essen, 1991; Jones e Powell, 1970). Si assumeva che l'elaborazione multisensoriale fosse confinata a specifiche aree nei cosiddetti stadi tardivi gerarchici, ipotesi corroborata dal fatto che queste aree terminano in aree corticali "associative", in altro modo, aree neocorticali associative 'di alto livello'. In altre parole, si pensava che l'integrazione sensoriale si verificasse ad uno stadio successivo all'elaborazione unisensoriale e deputata in zone della corteccia cosiddette polisensoriali o associative (Felleman e Van Essen, 1991).
Lo studio di Stein e Meredith (1983) sull'attivazione di neuroni nel collicolo superiore nella stimolazione multisensoriale, ha messo in crisi questa idea, dimostrando che aree pensate come esclusivamente multisensoriali sono coinvolte nell'elaborazione sensoriale di basso livello, e che possa riguardare processi additivi, superadditivi o subadditivi rispetto alle corrispettive risposte uni sensoriali (Budd e Timora, 2013). Più in generale, le risposte multisensoriali sono generalmente maggiori quando gli stimoli sensoriali 'unisensoriali' condividono determinate caratteristiche comuni, es. spaziali e temporali (Meredith et al, 1987; Meredith e Stein, 1996). Secondo Wallace et al. (1993) la risposta multisensoriale del collicolo superiore può riflettere l'influenza top-down nella convergere gli input dei neuroni corticali unisensoriali. Ad ogni modo, questo è solo uno dei primi studi di una lunga carrellata che mostrano che l'integrazione multisensoriale può verificarsi ad una fase precoce dell'elaborazione sensoriale in regioni corticali tradizionalmente pensate come esclusivamente uni sensoriali (per review vedere Schroeder e Foxe, 2005; Ghazanfar e Schroeder, 2006; Calvert, 2001).

Negli ultimi 25 anni sempre più studi tentano di analizzare come l'informazione delle diverse modalità sensoriali viene combinata dal cervello ed influire nel nostro comportamento e percezione. L'estensione dei meccanismi di integrazione multisensoriale dalle aree di convergenza multimodale alle aree primarie e dunque a tutti i livelli di elaborazione corticale ha fatto sì che venga proposto un ripensamento di tutti quei meccanismi cognitivi fino a non molto tempo fa modellizzati in una prospettiva essenzialmente unisensoriale. Alcuni studi mostrano i vantaggi comportamentali e percettivi di questa unione delle informazioni provenienti dalle diverse modalità sensoriali (Calvert, Spence e Stein, 2004; King e Calvert, 2001; Stein and Meredith, 1993; Stein et al., 2002a).
Diversi studi hanno corroborato la presenza e l'organizzazione di un numero di strutture corticali e sottocorticali in cui converge l'informazione di diverse modalità sensoriali, e che

l'integrazione neurale accompagna questa convergenza (Amedi et al.,2005; Beauchamp, 2005; Calvert et al., 1999; Calvert, Campbell e Brammer, 2000; Laurienti et al.,2002,2003; Lloyd et al.,2003; Martuzzi et al.,2007; Murray et al.,2005; Powers,Hevey eWallace,2012; Stevenson,Kim e James,2009).

Recentemente ha preso il via il dibattito se l'intera corteccia cerebrale (e per estensione l'intero cervello) possa essere considerata "multisensoriale" (Driver e Noesselt, 2008; Schroeder e Foxe, 2005; David Alais, 2010).

Uno dei risultati più importanti di questi studi è stato il riscontro che le caratteristiche degli stimoli che vengono combinati sono determinanti importanti del prodotto finale dell'interazione multisensoriale: primi tra tutti ad essere studiati, a livello del neurone individuale sono stati i fattori degli stimoli includono caratteristiche di spazio, tempo ed efficacia. Riguardo al tempo e allo spazio, è stato riscontrato che gli stimoli multisensoriali che sono spazialmente e temporalmente vicini tipicamente implicano un incremento della risposta neuronale (-in audiovisivo: Meredith, Nemitz, & Stein, 1987; Meredith et al., 1992; Meredith & Stein, 1986a, 1996; Royal et al., 2009;Wallace et al., 1996, 2004). Inoltre, stimoli che sono poco efficaci se presentati da soli mostrano un incremento della risposta proporzionale alla loro combinazione (James, Stevenson, & Kim, 2012; Meredith & Stein, 1983; Stein, Stanford, & Rowland, 2009). Questi principi integrativi di base rendono l'idea di un senso intuitivo che lo spazio e il tempo sono degli indicatori statistici potenti della probabilità che gli stimoli appaiano/si verifichino da uno stesso evento e che l'alta salienza o l'efficacia dello stimolo in una modalità ha bisogno di una piccola amplificazione. Studi recenti hanno ampliato il ruolo che questi fattori giocano nelle interazioni multisensoriali sottolineando la loro interdipendenza (Carriere, Royal, e Wallace, 2008; Ghose e Wallace, 2014; Krueger et al., 2009; Royal et al., 2009; Sarko et al., 2012; Sarko, Ghose, e Wallace, 2013). In quest'ottica, spazio tempo e efficacia non possono essere considerate entità indipendenti dal momento che la manipolazione di una, influisce sugli altri. Ad esempio la localizzazione spaziale influisce sull'efficacia degli stimoli e sui pattern di *firing* temporali associati. Questa ipotesi è stata corroborata sia da studi sulla facilitazione comportamentale e percettiva, per esempio degli stimoli che sono vicini spazialmente e temporalmente (es. Dixon & Spitz, 1980; Powers, Hillock, & Wallace, 2009), e la proporzione di benefici nella combinazione degli stimoli nelle diverse modalità appare essere più grande quando gli stimoli individuali sono poco efficaci (Diederich e Colonius, 2004; Kim e James, 2010; Kim et al., 2012; Nath e Beauchamp, 2011). Alcuni studi hanno illustrato da un punto di vista neuronale l'interdipendenza di questi principi nella performance e percezione umana (Macaluso et al., 2004; Royal et al., 2009; Stevenson et al., 2012c).

Per quanto riguarda gli aspetti temporali in cui si verifica il *binding* multisensoriale, vi sono diverse evidenze (vedi review di Budd TW e Timora JR, 2013) che dimostrano che l'integrazione multisensoriale può verificarsi ad una fase precoce dell'elaborazione sensoriale in regioni corticali tradizionalmente pensate come esclusivamente uni sensoriali. Se si considera che l'attività oscillatoria dei neuroni riflette una fluttuazione rapida e sincronizzata di un insieme di neuroni che alterna tra stadi alti e bassi di eccitabilità (Freeman, 1975), ne consegue che l'integrazione multisensoriale avviene tramite questa attività oscillatoria sincronizzata di gruppi neuronali (Engel et al, 2012). Sia che si parta da un'affermazione di '*binding* per sincronia' (Nozaradan et al., 2012) o di 'integrazione tramite la coerenza' (Senkowski et al., 2008), l'ipotesi di correlazione temporale afferma che la coerenza crossmodale delle oscillazioni crossmodali possono fornire i meccanismi necessari per la flessibilità e l'indipendenza dal contesto per il *binding* di informazioni provenienti dalle diverse modalità sensoriali.

Ursino e coll. (2014) esaminano diversi modelli: dal bayesiano, basato sulle probabilità come estimatore ottimale di cues esterni, al modello ispirato dalla biologia che prevede il collicolo superiore sia a livello di singolo neurone che come network di neuroni con l'enfasi negli schemi architettonici e meccanismi fisiologici, a modelli di plasticità sinaptica che riproducono lo sviluppo delle capacità integrative via *Hebbian-learning rules* o mappe *self-organizing*; e modelli di memoria semantica che implementano il significato degli oggetti come fusione di caratteristiche senso-motorie (*embodied cognition*). Gli autori sintetizzano alla fine i risultati presenti in letteratura che corroborano l'ipotesi circa il fatto che l'integrazione multisensoriale opera a diversi livelli del cervello: strutture sottocorticali (specialmente il Collicolo Superiore), cortecce associative di alto livello (es. regioni parietali posteriori) e perfino in aree corticali precoci (come le cortecce primarie) tradizionalmente considerate come coinvolte nei processi puramente unisensoriali.

Dibattito sulla modalità dominante.

Un argomento rilevante qualora ci si interroghi sull'interazione tra le modalità sensoriali è quello della dominanza tra modalità, soprattutto in casi di discordanza tra informazioni provenienti da canali sensoriali differenti. Al riguardo vi sono teorie diverse. Secondo l'ipotesi dell'attenzione

diretta, risulterà dominante la modalità sensoriale 'attesa' (Welch e Warren, 1980); tale ipotesi è poi statamodificata nell'ipotesi dell'appropriatezza, per cui il peso di una modalità nell'interazione con le altre dipende non dalle caratteristiche delle fonti d'informazione ma dal

grado di adeguatezza o appropriatezza della modalità in quel dato compito (Welch, 1999). Per l'ipotesi dell'attendibilità, ad essere invece dominante in un compito cognitivo sarebbe la modalità le cui stimolazioni apportano, per quel compito, le informazioni maggiormente pregnanti (Schwartz et al., 1998). Infine, secondo l'ipotesi della discontinuità, risulta dominante la modalità in cui la stimolazione è maggiormente discontinua (Shams et al., 2002). Andersen e colleghi tuttavia offrono considerazioni teoriche a favore dell'idea che queste ipotesi dovrebbero essere considerate come fattori che contribuiscono alla dominanza di ciascuna modalità e non come condizioni tutto o nulla (Andersen et al., 2004).

Ernst, Banks, 2002 propongono un principio generale per determinare il grado di dominanza relativa di una modalità sull'altra (nei casi in cui la discprepanza non è larga al punto da far rifiutare una delle due fonti di informazione), detto "*Maximum Likelihood Estimate*", ovvero un metodo statistico che il sistema nervoso adotterebbe per creare una stima ottimale, con una varianza minima dello stato del mondo date e proprie informazioni. L'autore conclude affermando che le modalità sensoriali che partecipano di una medesima unità si influenzano, in misura diversa; comunque, anche nel caso di dominanza dell'una sull'altra la modalità non dominante non è semplicemente ignorata, ma influenza a sua volta quella che domina, di modo che il sistema percettivo sembra non ignorare nessuna delle informazioni in suo possesso.

Inoltre, diversi lavori descrivono l'influenza modulatoria che una modalità 'non dominante' può avere sull'elaborazione della modalità 'dominante'(in audiovisivo- Hackett & Schroeder, 2009; Zion Golumbic et al., 2013).

Secondo Spence (2010) la visione sembra essere dominante nei compiti spaziali, ma è possibile l'influenza degli stimoli uditivi e tattili che influenzano la visione e l'elaborazione dell'informazione visiva. La dominanza dell'uditivo sul visivo sembra prevalere in compiti temporali (es. illusione dei due flash). La presentazione simultanea di multipli stimoli uditivi o tattili influenza l'attività neurale a livelli precoci dell'elaborazione dell'informazione, inclusa la corteccia visiva primaria.

1.2.3 *Binding* audio-tattile: aspetti generali.

Finora si è parlato di *binding*, o meglio i problemi del *binding*, in generale. In questo paragrafo si farà riferimento ad un tipo specifico di *binding* crossmodale, il *binding* audio-tattile o audio-aptico (a seconda se si vorrà sottolineare l'aspetto tattile o cinestesico del fenomeno).

Come accennato nella parte iniziale, queste due modalità presentano diverse similarità, sia da un punto di vista fisiologico (recettori con forze meccaniche), che neuroanatomico (adiacenza neuroanatomica), e funzionale (elaborazione dell'informazione in modo sequenziale).
George A. Gescheider sin dal 1965 ha studiato degli effetti di interferenza e influenza del tatto sul suono e viceversa, fornendo le basi per lo studio del *binding* audio-tattile.
Ma vediamo cosa riporta la letteratura nello specifico circa l'influenza reciproca di una modalità sull'altra e la loro integrazione a livello percettivo, prendendo in esame sia gli studi comportamentali e psicofisici, sia studi sui correlati neurali.

***Binding* audio-tattile: studi comportamentali e psicofisici.**

Influenza del suono sulla percezione tattile.

In letteratura, sono presenti innumerevoli studi che dimostrano l'influenza della modalità uditiva sulla modalità tattile, sia in termini di facilitazione, che di mascheramento. Di seguito, riporto alcuni tra i più rilevanti esempi, in ordine cronologico.
Gescheider dimostrò per primo che il suono può influenzare, e nello specifico, facilitare, la percezione di eventi tattili, per esempio aumentare la soglia tattile assoluta e fenomeni di mascheramento (per esempio, si vedano gli studi del 1970 e del 1974) ed un effetto facilitatore per la risoluzione temporale tra eventi tattili (per esempio, si veda lo studio del 1967). In particolare, Gescheider e coll (1974) hanno verificato che l'elaborazione uditiva può avere un effetto inibitore sull'elaborazione tattile utilizzando una procedura di mascheramento per valutare la soglia di detezione dei pattern vibrotattili alla punta del dito in base ad un suono presentato simultaneamente. Ad ogni modo, si è visto che si ha un abbassamento della soglia di detezione degli stimoli tattili vibratori quando è presentato in simultanea ai suoni se il contenuto della frequenza è uguale in termini di contenuto (REF) mentre in altre occasioni il suono produce un effetto di smascheramento (REF) (si noti però che tra gli esperimenti riportati vi sono differenze metodologiche (REF).
Uno studio ben noto in letteratura sull'influenza degli stimoli uditivi sulla percezione tattile, è il fenomeno riportato da Jousmäki e Hari (1998). Gli autori riportano un fenomeno chiamato "*parchment-skin illusion*", in cui, si potrebbe dire, c'è un cambiamento del 'contenuto' della frequenza (cit. da Soto-Faraco e Deco, 2009): nel momento in cui determinati suoni sono sincroni allo strofinio delle mani, la sensazione tattile viene alterata. I soggetti ascoltano il rumore prodotto dallo strofinio delle mani, che viene attenuato o amplificato (di +o-15 dB, per un massimo di 20-40 dB). I soggetti riportano delle differenze di sensazione tra maggiore ruvidità associata a umidità e minore ruvidezza associata ad una sensazione di asciutto. In altre

parole, ad alte frequenze la pelle viene percepita simile alla carta. In un secondo esperimento, gli autori riscontrano che con un ritardo di 100 ms del feedback uditivo, l'effetto non si verifica. Gli autori interpretano questi dati concludendo che il *binding* multisensoriale efficace deve verificarsi entro una certa finestra temporale ed ipotizzano che questo effetto rifletta un'integrazione intersensoriale onnipresente che potrebbe fornire informazioni in merito a caratteristiche quali la ruvidità o rigidità di tessuti o materiali.

Guest e coll. (2002) riprendono lo studio di questo fenomeno con compito di scelta forzata con stimoli di diversa ruvidezza. Anche essi riscontrano che la percezione della ruvidezza tattile è modulata dalla frequenza degli stimoli uditivi. In particolare, gli autori riscontrano che abbassando le alte frequenza aumenta la levigatezza di una superficie. Inoltre, in un secondo compito simile a quello svolto da Jousmäki e Hari (1998) vengono replicati i dati circa la percezione delle proprie mani come più o meno umide in base al feedback uditivo (vedi esperimento su). Infine, in un terzo esperimento gli autori mostrano come aumentando l'intervallo tra il feedback uditivo e lo strofinio delle mani viene ridotto l'effetto. Gli autori concludono dicendo che gli stimoli uditivi possono variare la salienza della percezione tattile della superficie.

Nel 2003 Zampini e coll. hanno studiato l'influenza del suono prodotto da un spazzolino elettrico sulla percezione dello stesso. I partecipanti dovevano usare lo spazzolino e giudicare la piacevolezza e rugosità della stimolazione vibrotattile che percepivano sui loro denti. I partecipanti hanno riportato una maggiore piacevolezza e minore ruvidezza quando il livello del suono era ridotto o le frequenze alte del suono venivano attenuate. La sensazione veniva quindi, modificata alterando le variazioni del feedback uditivo prodotto dall'azione dello spazzolamento.

Sempre con la stessa ipotesi di ricerca, nel 2004 Zampini e Spence indagano l'influenza del suono nel giudicare la croccantezza e freschezza di patatine fritte. I partecipanti dovevano dare un giudizio di freschezza e croccantezza di una patatina mordendola con i denti frontali e in simultanea veniva presentato un suono prodotto durante una masticazione. I partecipanti hanno giudicato la freschezza e croccantezza secondo il volume e frequenza del segnale uditivo presentato: la patatina veniva giudicata come più croccante quando il livello del suono veniva incrementato o venivano selettivamente amplificate le frequenze più alte (in un range di 2 kHz-20 kHz).

Nel 2005 Bresciani e coll. Hanno riscontrato che sequenze di 'beep' uditivi possono modificare la percezione di sequenze di 'picchiettii' tattili fatti sulla punta dell'indice (sequenze da due a quattro picchettii). In un primo esperimento viene chiesto ai partecipanti di focalizzarsi sulle

sequenze tattili (presentazione simultanea delle sequenze tattili e ‘beep’ uditivi, congruenti o incongruenti per numero), ma gli autori hanno riscontrato che la percezione degli stimoli tattili veniva sistematicamente modulata dal numero dei beep, quando le due stimolazioni erano simili. In un secondo esperimento gli autori inseriscono un bias, per cui le due sequenze erano presentate in modo asincrono. L’influenza dei beep uditivi si verificava solo quando erano sincronizzati agli stimoli tattili, ragion per cui gli autori concludono che il sistema nervoso centrale integra questi due stimoli solo quando li reputa come ‘generati’ dallo stesso stimolo (quando sono appunto, in sincronia).

Questi e studi simili, vengono riportati nella rassegna bibliografica svolta da Spence e Zampini nel 2006, nella quale vengono riportati diversi studi in cui stimoli uditivi associati al contatto e interazione con superfici differenti (per esempio carta vetrata abrasiva) e prodotti (il già citato spazzolino da denti, ma anche spray dell’aereosol, macchine) modificano la percezione degli oggetti, senza che ne siamo consapevoli. L’autore suggerisce che ciò potrebbe comportare il cambiamento della percezione della qualità ed efficienza di tali prodotti, senza contare gli innumerevoli sviluppi sulla percezione anche ad un livello semantico.

Suzuki e coll. (2008) effetti del suono sulla percezione tattile della rugosità e lunghezza. I risultati mostrano che il rumore bianco (suono complesso) diminuisce significativamente l'inclinazione della funzione della stima di ruvidezza, mentre non influisce sulla funzione della stima della lunghezza, per cui suoni complessi influenzano la percezione tattile della ruvidità anche quando essi sono apparentemente irrilevanti per l'esplorazione delle superfici. Gli autori interpretano questi dati come indice che l'elaborazione dei suoni complessi possa essere correlata all'elaborazione della ruvidità tattile, mentre è indipendente dall'elaborazione della lunghezza tattile.

Ro e coll. (2009) studiano come gli stimoli uditivi irrilevanti per un compito alterano la detezione di stimoli somatosensoriali vicini alla soglia. Gli autori svolgono tre esperimenti in cui mostrano che gli stimoli uditivi simultanei incrementano la sensibilità ma non il biases delle risposte per la detezione di stimoli cutanei elettrici sulla mano (esperimento 1), che l'aumento della percezione somatosensoriale è spazialmente specifico -solo suoni monoaurali sullo stesso lato incrementano la detezione- (esperimento 2), che gli effetti dell'udito sulla percezione tattile dipendono dalla frequenza-solo suoni con la stessa frequenza di quella vibrotattile incrementano la detezione tattile (esperimento 3). Gli autori interpretano questi risultati come indice che l'informazione uditiva influenza la percezione tattile in modo altamente sistematico e suggerisce che simili meccanismi di codifica possono essere alla base dell'elaborazione dell'informazione di queste diverse modalità sensoriali.

Influenza del tatto sulla percezione del suono.

In letteratura, non mancano anche casi in cui è stata studiata l'influenza della stimolazione tattile sulla percezione uditiva. Di seguito, riporto alcuni tra gli studi più rilevanti in merito.

Nel 1932 Paul von Schiller notò per primo che alcuni suoni (scoppi di rumore o toni ripetuti ad intervalli regolari) possono influenzare la percezione della ruvidità.

Nonostante ciò, lo studio mirato ad indagare l'influenza del tatto sulla percezione uditiva, è più di recente rispetto allo studio dell'influenza del suono sulla percezione tattile.

Nel 2002, Caclin e coll. hanno svolto 4 esperimenti per studiare l'influenza del tatto in compiti di localizzazione uditiva cercando di riprodurre una sorta di effetto ventriloquo audio-tattile (Bertelson, 1998). I partecipanti dovevano discriminare se i suoni venivano presentati alla loro destra o sinistra. Gli stimoli uditivi venivano presentati da soli, o in concomitanza di una stimolazione tattile alla punta delle dita. Nei primi tre esperimenti, gli autori concludono che il giudizio della localizzazione dei suoni veniva influenzato dalla stimolazione tattile, quando avveniva in modo sincronizzato alla stimolazione uditiva, rispetto a quando le due stimolazioni non erano sincronizzate. Dal momento che questo lasciava aperta la domanda se il tatto influenza la localizzazione uditiva perché era distraente, e non perchè avvenisse un'effettiva interazione audio-tattile nella localizzazione, svolgono il quarto esperimento in cui viene aggiunta una condizione di stimolazione tattile centrale. In questo caso, gli autori indicano che viene escluso l'effetto di un'attenzione endogena (o volontaria) e che il tatto influenza la localizzazione sonora un pò come per l'effetto ventriloquo. Dal momento che il prestare più attenzione alla modalità tattile non ha incrementato il bias di localizzazione sonora nella condizione sincrona (e quindi il fenomeno era indipendente rispetto ad un volontario direzionamento dell'attenzione), concludono che ciò mostra una influenza del tattile sulla modalità uditiva a prescindere dall'orientamento attentivo.

Nel 2004 Schürmann e coll. svolgono un esperimento psicofisico in cui i soggetti svolgono un compito di comparazione del volume di suoni (devono equiparare un suono partendo dopo aver ascoltato un altro suono) mentre toccano o no un tubo che vibra. I soggetti hanno scelto intensità più basse del 12% quando toccano il tubo. Gli autori interpretano questo dato come un effetto di facilitazione dell'interazione audio-tattile.

Salvador Soto-Faraco e coll. (2004) esaminano gli effetti di un flusso di movimento apparente in una modalità irrilevante (uditiva o tattile) sulla percezione di un flusso di movimento apparente in un'altra modalità (tatto o udito, rispettivamente). Gli autori riscontrano un effetto di congruenza (bassa prestazione quando la direzione del movimento della modalità irrilevante

è incongruente con la direzione del target) per le due possibili combinazioni. Inoltre, gli autori indicano che questo effetto di congruenza è asimmetrico: i distruttori di movimento tattile hanno un'influenza maggiore sulla percezione del movimento uditivo, rispetto che il viceversa. In un secondo esperimento gli autori verificano l'influenza della posizione delle braccia (incrociate o no) nella percezione del movimento apparente e riscontrano che quando i soggetti hanno le braccia incrociate l'effetto è inferiore. Gli autori interpretano i risultati alla luce della rappresentazione dello spazio per la modalità tattile e uditiva.

Similmente, nel 2005 Sanabria e coll. studiano l'influenza della posizione delle mani sull'interazione audio-tattile spaziotemporale svolgendo due esperimenti. Nel primo esperimento i partecipanti devono indicare la direzione di un flusso di suoni (uno suono in una orecchia, e un secondo suono nell'altra orecchia) e ignorare una stimolazione tattile concomitante (che poteva essere spazio-temporalmente congruente o incongruente a quella uditiva). Quando i partecipanti avevano le mani non incrociate era presente un decremento significativo della performance per le prove incongruenti, rispetto a quando incrociavano le mani. Nel secondo esperimento invertono la modalità target di riposta e ritrovano lo stesso effetto, concludendo circa l'importanza della postura corporea nel rispondere al compito e sulle interazioni crossmodali più in generale.

Gillmeister ed Eimer (2007) studiano gli effetti della modalità tattile su quella uditiva. Gli autori trovano che segnali tattili irrilevanti facilitano la detezione di toni deboli, ed aumentano la stima della loro intensità. Questi effetti di facilitazione crossmodale vengono riscontrati quando si confronta la stimolazione audio-tattile sincrona con quella asincrona, e sono più evidenti per i suoni più deboli rispetto a quelli più forti. Gli autori interpretano questi risultati in termini di meccanismi di integrazione multisensoriale che aumenta la forza dei segnali uditivi e aderisce ai ruoli di efficacia inversa e della co-occorrenza temporale (ma non spaziale). Gli autori ipotizzano che questa integrazione potrebbe essere mediata dai neuroni multisensoriali audio-tattili in regioni della corteccia associativa uditiva che sono anche coinvolti nella detezione uditiva e discriminazione del volume.

Binding **audio-tattile: network neurali.**

Come per gli studi comportamentali, il *binding* audio-tattile è stato indagato dalle neuroimmagini. Verranno di seguito riportati solo studi umani, sebbene storicamente in una fase iniziale dello studio del *binding* audio-tattile sono presenti in letteratura una molteplicità di studi sui primati non umani.

Come accennato, a livello neuroanatomico, l'area primaria uditiva (A1) è collocata nel piano temporale superiore che è adiacente alla corteccia somatosensoriale secondaria nell'opercolo parietale (si veda il caso di sinestesia audio-tattile riportato da Ro et al., 2007). Questo lascia supporre che, a livello neuroanatomico la loro interazione sia plausibile.

Storicamente, partendo da studi con animali e poi umani, si è riscontrata l'attivazione della corteccia uditiva in risposta a stimoli tattili (esempi Foxe et al. 2002; Kayser et al. 2005; Schurmann et al. 2006) , e successivamente anche il viceversa, ovvero anche la corteccia somatosensoriale era reattiva a stimolazione uditiva (Ro et al., 2013).

Di seguito verranno esposti gli studi in ordine cronologico.

Nel 2002 Foxe e coll. tramite tecnica fMRI trovano che gli input uditivi e somatosensoriali convergono in una subregione della corteccia uditiva lungo il giro temporale superiore. Inoltre, la stimolazione simultanea in entrambe le modalità sensoriali risulta in un aumento dell'attività così come predetto dalla somma delle risposte degli input unisensoriali, dimostrando quindi che vi sia una regione di convergenza per l'integrazione multisensoriale (un'area simile è stata riscontrata precedentemente negli studi con animali). Secondo gli autori questa convergenza uditiva e somatosensoriale nella corteccia uditiva precoce contribuisce a corroborare l'ipotesi che l'integrazione multisensoriale precoce nell'elaborazione corticale gerarchica coinvolge aree cerebrali precedentemente assunte come esclusivamente unisensoriali.

Successivamente, sono stati svolti studi in cui si è analizzata più specificatamente, l'interazione audio-tattile a livello neurale.

In uno studio successivo, Foxe e coll. (REF) confrontano l'attivazione delle aree durante stimolazione unisensoriale (uditiva o tattile) e multisensoriale (audio-tattile). Da notare è che gli stimoli uditivi usati sono stati il suono emesso dalla carta vetrata o da una superficie ruvida (simile ai suoni di feedback dell'effetto "*parchment-skin illusion*" (Jousmäki e Hari, 1998) e come stimolazione tattile i soggetti dovevano fare rotolare un rullo di legno tra le dita indice e medio della mano destra. I risultati mostrano che gli input uditivo e somatosensoriale convergono in una subregione della corteccia uditiva lungo il giro temporale superiore. Inoltre, la stimolazione simultanea di entrambe le modalità sensoriali mostra un'attività in eccesso così come predetta dalla somma delle risposte degli input uni sensoriali, dimostrando l'integrazione multisensoriale in questa regione di convergenza. Gli autori concludono dicendo che l'integrazione uditiva e somatosensoriale avviene in aree deputate all'elaborazione uditiva precoce, confermando l'ipotesi circa l'elaborazione multisensoriale implica l'integrazione a stadi precoci dell'elaborazione corticale.

Lütkenhöner e coll. (2002) utilizzano la tecnica MEG con lo stesso fine. Anche loro presentano stimolazione uni sensoriale (toni binaurali a 1047-Hz e 60 dB di sensazione e come stimolo tattile la pressioni pulsate al pollice destro) e audio-tattile simultanea. I risultati mostrano una chiara interazione audio tattile nell'emisfero controlaterale alla stimolazione tattile per la maggior parte dei partecipanti (in sei partecipati su otto, tre soggetti hanno mostrato l'interazione nell'emisfero ipsilaterale). Gli autori suggeriscono che andrebbero svolti ulteriori studi per indagare se l'interazione audio tattile riscontrata nell'attivazione di SI sia correlata con la localizzazione spaziale degli stimoli (informazione che secondo gli autori viene comunque veicolata da questo tipo di interazione, es. qualcuno bussa alla porta veicola informazioni spazio-temporali). Inoltre, ipotizzano che la polarità di questa interazione corrobori l'ipotesi che lo stimolo uditivo risulta in una parziale inibizione in SII. In due soggetti si nota un forte indicatore del contributo uditivo nell'interazione (sebbene in esmisferi diversi).
Un altro studio MEG è stato svolto da Gobbelé e coll. (2003). Viene registrata la stimolazione uni sensoriale (uditiva: toni presentati ad un orecchio; tattile: impulsi elettrici per un lato del nervo mediano del polso) e audio-tattile simultanea, alternate sul lato destro e sinistro. L'interazione è risultata maggiore nell'emisfero sinistro rispetto al destro. Inoltre, nella maggior parte dei soggetti, le risposte nella condizione di stimolazione simultanea audio-tattile risultavano simili alla stimolazione tattile e non alla stimolazione uditiva. Gli autori interpretano questo dato come suggerimento circa il fatto che vi sia una soppressione dell'elaborazione uditiva durante la concordanza spazio-temporale audio tattile nella quale la componente tattile sembrerebbe essere in questi casi, più saliente.
Entrambi questi studi (Lütkenhöner e coll., 2002 e Gobbele´ e coll., 2003) corroborano l'ipotesi circa il fatto che l'attivazione legata a stimolazioni audio tattili correli con la salienza degli stimoli utilizzati. Infatti, nello studio di Lütkenhöner e coll. (2002) quando i suoni sono più salienti degli stimoli tattili (per esempio toni binaurali di 1-kHz combinati con pressioni sul pollice) è presente una inibizione parziale nella corteccia somatosensoriale secondaria (SII) e in Gobbele´ e coll. (2003) la salienza tattile maggiore rispetto a quella degli stimoli uditivi sembra legata ad una soppressione dell'elaborazione uditiva quando gli stimoli sono spazialmente e temporalmente concordanti.
Altro studio MEG di Caetano e Jousmäki (2006) ha come scopo quello di indagare se la stimolazione vibro tattile (vibrazioni di un tubo tenuto con la mano a 200-Hz mediamente oltre i 19.5 dB della soglia percettiva tattile soggettiva) attiva le aree corticali uditive. Tutti i soggetti hanno riportato di percepire un suono mentre toccavano il tubo vibrante. La stimolazione vibro tattile elicita un'attivazione in aree specifiche in quasi tutti i soggetti (dieci su undici), mentre

non si riscontra nessun tipo di risposta da parte di queste aree in assenza di stimolazione vibro tattile. Gli autori interpretano questi risultati come prova che gli stimoli vibrotattili implicano l'attivazione della corteccia uditiva in soggetti udenti (si rifanno a precedenti studi con soggetti sordi). Gli autori concludono dicendo che la convergenza degli input vibrotattili nella corteccia uditiva riscontrata in soggetti udenti adulti conferma dati precedenti con soggetti sordi.

Nel 2006 Shuermann e coll. indagano le aree uditive che potrebbero essere attivate anche con stimoli tattili (vibrazioni e pulsazioni tattili a 200Hz sulla punta delle dita e sui palmi, mentre per il suono è presentato del rumore). Gli autori riscontrano una co-attivazione audio-tattile in una regione della cintura uditiva posteriore (85-mm³). Gli autori interpretano questi risultati affermando questa co-attivazione correla con l'ascolto facilitato ad un livello comportamentale, riflettendo l'analisi temporale come per la vibrazione sonora. Ad ogni modo, visto che anche le pulsazioni senza vibrazione attivano parti della cintura posteriore uditiva, gli autori affermano che quest'area possa essere legata all'elaborazione degli eventi audiotattili che sorgono durante il contatto dinamico delle mani con l'ambiente. Gli autori riportano inoltre che negli esseri umani, studi cerebrali non invasivi hanno rilevato aree neurali correlate all'interazione audio tattile, coinvolte in modo differente in base alla salienza e alla qualità degli stimoli: corteccia somatosensoriale secondaria (SII), area della cintura uditiva e la corteccia parietale posteriore (si vedano i già descritti Foxe e coll. 2002; Lütkenhöner e coll., 2002; Gobbele´ et al., 2003; Caetano and Jousmäki, 2006).

Nel 2009 Reiner e coll. hanno come obiettivo quello di indagare se le due vie uditiva e tattile convergono o meno in un'area cerebrale multisensoriale specializzata. Tramite fMRI, gli autori comparano negli stessi soggetti, l'attivazione cerebrale riferita alla localizzazione e identificazione di stimoli comparabili uditivi e vibrotattili. I risultati indicano che il giro frontale inferiore destro (IFG) e insula destra e sinistra sono più attivate durante le condizioni di localizzazione rispetto alla localizzazione in entrambe le modalità. Viene riscontrata una dissociazione inversa, ovvero un'attivazione maggiore dei lobi parietali inferiori destro e sinistro (IPL), il lobo parietale superiore sinistro (SPL) e il precuneo destro del SPL durante la condizione di localizzazione per entrambe le modalità. Gli autori concludono proponendo che le aree specializzate nel IFG destro e nell'insula destra e sinistra sono operatori multisensoriali per l'elaborazione dell'identità degli stimoli, mentre parte del IPL destro e sinistro e il SPL sono specializzati nell'elaborazione degli attributi spaziali indipendentemente dalla modalità sensoriale.

Bolognini e coll. (2009) studiando il tatto hanno riscontrato attivazioni di aree uditive. Essi infatti partono dalle evidenze che sostengono che l'attività neurale dell'integrazione

multisensoriale coinvolge anche gli stadi di elaborazione precoce dell'info sensoriale. In particolare essi indagano usando la TMS se il giro temporale superiore (un'area specifica per la modalità uditiva) è anche critica per l'elaborazione di eventi tattili (stimolano il giro temporale superiore sinistro e la corteccia somatosensoriale primaria sinistra durante lo svolgimento di compiti tattili di discriminazione temporale e spaziale). La stimolazione del giro temporale superiore interferisce con i compiti di elaborazione tattile temporale (limitato al tatto contro laterale e dovuto ad ridotta sensibilità percettiva; l'interferenza spaziale si verificava con una stimolazione SI a 60–120 msec). Gli autori interpretano i risultati come l'implicazione causale delle aree uditive nell'elaborazione della durata degli eventi somatosensoriali, suggerendo che il giro temporale superiore possa giocare un ruolo sopramodale nella percezione temporale, oltre che corroborare l'ipotesi di partenza circa il fatto che l'integrazione multisensoriale avviene ad uno stato precoce dell'elaborazione dell'informazione (per dettagli si veda paragrafo sulla temporalità).

Kim e Zatorre (2011) hanno indagato il ruolo del complesso laterale occipitale (LOC) nell'elaborazione della forma esaminando (utilizzano tecnica fMRI) l'attività neurale associata con l'apprendimento dell'informazione uditiva codificata insieme ad una forma tattile (lo stimolo tattile poteva essere coerente o no con il suono ascoltato). Gli autori riscontrano migliore performance dopo un training, mentre l'attivazione di LOC era attivata in ugual modo durante le condizioni tattili e uditiva sia prima che dopo una fase di apprendimento (*training*), non mostrando nessun cambiamento significativo nella sua attivazione; mentre la corteccia uditiva e il LOC mostrano connettività funzionale più forte dopo il *training*. Gli autori attribuiscono questi risultati come prova a sostegno che il LOC è disponibile a sistemi sensoriali differenti per l'elaborazione della forma e che il training alla sostituzione della sensazione audio-tattile comporta cambiamenti neurali permettendo un accesso più diretto ed efficiente a questi siti tramite il sistema uditivo.

Kassuba e coll. (2013) partono dal proposito di identificare i correlati neurali del *binding* audio-aptico di caratteristiche di oggetti. I soggetti, dentro la fMRI, devono abbinare un oggetto target ad un oggetto campione attraverso la modalità uditiva e tattile. Introducendo un ritardo tra la presentazione dei due oggetti, gli autori si prefiggono di indagare le aree di attivazione in compiti in cui è presente un ritardo tra gli stimoli e quindi poter valutare il match dell'uditivo sull'aptico e dell'aptico sull'uditivo. Inoltre, gli autori ipotizzano che solamente le caratteristiche semanticamente coerenti tra le caratteristiche uditive e aptiche degli oggetti attivano aree corticali adibite alla rappresentazione concettuale degli oggetti. I risultati mostrano che il giro fusiforme sinistro (FG) e il solco temporale superiore posteriore (pSTS)

presentano un incremento di attivazione durante l'appaiamento cross-modale degli stimoli semanticamente congruenti, ma non degli stimoli incongruenti. Nel FG questo effetto è stato riscontrato in entrambe le direzione (aptico sull'uditivo e viceversa), mentre per il pSTS si verifica un effetto *matching* cross-modale solo per i target uditivi congruenti. Le cortecce associative uditiva e somatosensoriale mostrano un aumento di attività durante l'appaiamento di oggetti cross-modali, che risulta essere indipendente dalla congruenza semantica. Gli autori interpretano questi due risultati come prova dell'interazione multisensoriale a stadi gerarchici differenti dell'elaborazione uditiva e aptica degli oggetti. Le interazioni cross-modali oggetto-specifiche culminano nel FG sinistro, il che provvedere ad una zona di convergenza di ordine superiore per la conoscenza concettuale degli oggetti. Gli autori concludono proponendo che tatto e udito interagiscono a vari livelli dell'elaborazione degli oggetti con modulazioni cross-modali a livello precoce delle cortecce associative e a livello intermedio nelle regioni di riconoscimento degli oggetti modalità-sensazione (pSTS e verosimilmente LO) e converge in rappresentazioni degli oggetti concettuali di alto livello nel FG sinistro. Propongono per il futuro un'analisi del ruolo del FG nel riconoscimento di oggetti multisensoriali ed in modo particolare nell'integrazione dell'informazione degli oggetti uditivi ed aptici.

Riguardo alla plasticità crossomodale audio-tattile conseguente a deprivazione sensoriale, alcuni studi riportano che, mentre si elaborano informazioni tattili, i soggetti sordi congeniti mostrano una maggiore attivazione delle aree considerate tradizionalmente come parti della corteccia uditiva (giro di Heschl e aree circostanti) (REF). Inoltre, è stata riscontrata una maggiore sensibilità tattile nei sordi congeniti (REF), presunta conseguenza della plasticità. Questo dato dà speranza alla teoria della sostituzione sensoriale per lo sviluppo di strumenti tattili per soggetti sordi (si veda capitolo sulla teoria della sostituzione sensoriale).

Binding audio-tattile: aspetti temporali.

In questa sezione vengono riportati studi che indagano i tempi di attivazione delle modalità uditiva e tattile quando stimolate in simultanea.

Foxe e coll. (2000) indagano il corso temporale (ERPs) e la topografia dello scalpo (scalp current density, SCD) dell'interazione multisensoriale conseguente alla stimolazione simultanea uditiva e somatosensoriale. I potenziali evento correlati (64 elettrodi) vengono registrati mentre vengono presentati stimoli solo uditivi (1000Hz), solo somatosensoriali (impulsi elettrici al nervo mediano) e stimolazione combinata. L'analisi spazio-temporale degli ERP e le topografie di densità corrente dello scalpo rivelano un'interazione uditiva-somatosensoriale sopra lo scalpo centrale/postcentrale che inizia approssimativamente 50 ms

dopo la presentazione dello stimolo. Entrambi i risultati (topografico e temporale) corroborano l'ipotesi di una integrazione multisensoriale precoce nella gerarchia dell'elaborazione corticale delle aree pensate originariamente come esclusivamente unisensoriali. In particolare, L'analisi spazio-temporale e la topografia della densità dello scalpo, rileva un effetto interazione positivo dopo 50 ms dopo la presentazione dello stimolo nella corteccia somatosensoriale nella regione del giro postcentrale e un'interazione tardiva a 70-80 ms nella regione delle cortecce uditive posteriori vicino al piano temporale superiore.Questo studio supporta al teoria dell'integrazione sensoriale gerarchica con un'elaborazione sensoriale precoce precedente e/o in aggiunta all'integrazione in aree eteromodali.

Nello studio di magnetoencefalografia, Lütkenhöner e coll. (2002) riportano che a livello temporale, la maggiore deflezione di polarità opposta viene raggiunta a 140 e 220 ms e la prima deflezione si ha nella regione della corteccia somatosensoriale secondaria (SII). Gli autori discutono questa differenza nell'attivazione emisferica ipsilaterale o controlaterale come indicativa di una differenza interindividuale nella percezione simultanea audio-tattile.

Nello studio di Gobbelé e coll. (2003), sempre con tecnica MEG, viene riportata una attivazione significativamente differente dalla somma algebrica delle risposte audio-tattili simultanee a 75–85 e 105–130 ms, indicando una interazione audio tattile soppressiva. Rilevano inoltre un'interazione precoce nella corteccia parietale posteriore controlaterale e un'interazione più tardiva nell'opercolo parietale controlaterale tra la corteccia SII e la corteccia uditiva. L'interazione è risultata maggiore nell'emisfero sinistro rispetto al destro. Come già menzionato, gli autori interpretano questo dato come suggerimento circa il fatto che vi sia una soppressione dell'elaborazione uditiva durante la concordanza spazio-temporale audio tattile nella quale la componente tattile sembrerebbe essere in questi casi, più saliente.

Nello studio di Caetano e Jousmäki (2006), in cui si è indagato se la stimolazione vibro tattile attivasse le aree corticali uditive (si veda paragrafo sulle aree neurali), i risultati mostrano una prima risposta alla vibrazione intorno ai 60 ms nella corteccia primaria somatosensoriale in tutti i soggetti. Successivamente, a 100- a cui segue l'attivazione nelle cortecce uditive bilateralmente (5 soggetti) o unilateralmente (5 soggetti) e un'attivazione nella corteccia SII, contro laterale (3 soggetti) o ipsilaterale (4 soggetti). In nove soggetti è stata riscontrata un'ulteriore attivazione uditiva sostenuta a 200-700 ms, bilateralmente (due soggetti) e ipsilateralmente (sette soggetti).

In uno studio con TMS (Transcranial magnetic stimulation), Bolognini e coll. (2006) hanno effettuato stimolazioni TMS al giro temporale superiore (STG), un'area considerata come specificatamente uditiva. La stimolazione del STG sinistro e alla corteccia somatosensoriale

primaria (SI) è stata effettuata a diversi intervalli temporali (60, 120 e 180 ms) e durante un compito di discriminazione temporale tattile (esperimento 1) e di discriminazione spaziale tattile (esperimento 2). L'elaborazione tattile è stata interrotta quando la TMS veniva applicata alla SI a 60 ms dopo la presentazione degli stimoli tattili, confermando la specificità di questa regione per l'elaborazione tattile. La stimolazione TMS sopra il STG influenzava anche l'elaborazione temporale tattile, ma con 180 ms di ritardo. In entrambi i casi, la stimolazione TMS inficiava l'elaborazione tattile quando era controlaterale e dovuta ad una sensibilità percettiva ridotta. Di contro, l'elaborazione tattile spaziale veniva inficiata solo per TMS alla SI a 60-120 ms. Dal momento che l'STG (area considerata come uditiva) si arriva per elaborare la durata di eventi tattili, gli autori concludono che quest'area possa avere un ruolo sovramodale nella percezione temporale (già accennato nel paragrafo precedente sulle basi neurali).

Scheider e coll. (2011) analizzano come l'esplorazione aptica di oggetti familiari influenzi l'elaborazione neurale del suono degli oggetti presentati successivamente. Gli autori utilizzano l'EEG per misurare l'attività gamma-banda (GBA) in un paradigma di *priming* da aptico a uditivo (ovvero gli stimoli aptici utilizzati come prime e suoni di oggetti come target). *Prime* e oggetti target potevano essere congruenti o incongruenti da un punto di vista semantico. I partecipanti devono classificare gli oggetti in base al suono. I risultati mostrano un *priming* aptico-uditivo associato ad un potere totale GBA (250-350 ms) per gli input congruenti semanticamente ed un effetto addizionale della congruenza semantica su GBA evocati (50-100 ms). Inoltre, la ricostruzione della sorgente del GBA totale rivela gli effetti della congruenza semantica nel lobo temporale laterale sinistro (interpretato dagli autori come possibile indicatore dell'abbinamento delle informazioni provenienti da diverse modalità), e per gli input semanticamente incongruenti la GBA totale è aumentata nelle cortecce frontali mediali (interpretato dagli autori come possibile indicatore dell'elaborazione o detezione dell'informazione conflittuale). Gli autori concludono affermando che il *priming* semantico dell'esplorazione degli oggetti in modo aptico influenza l'elaborazione degli input uditivi nel lobo temporale laterale e suggerisce un ruolo importante nell'attività oscillatoria dell'elaborazione multisensoriale.

Butler e coll. (2012) si interrogano se il sistema sensoriale calcola la frequenza (intesa come un costrutto percettivo multisensoriale) indipendentemente prima che questa informazione sia combinata o se questa sua caratteristica è calcolata in un modo integrato durante l'elaborazione sensoriale preattentiva. Essi utilizzano il *mismacht* negativo, la risposta elettrofisiologica che indica la detezione preattentiva di un contesto con stimolazione regolare, viene usata come misura dipendente. La registrazione elettrofisiologica ad alta densità viene svolta durante la

presentazione di stimoli somatosensoriali, uditivi e audio-somatosensoriali, standand o deviati (ovvero che differiscono per frequenza). Effetti multisensoriali vengono identificati a circa 200 ms, con il mismatch multisensoriale negativo (multisensory mismatch negativity (MMN) significativamente diverso dalla somma del MMNs uni sensoriali. Gli autori concludono affermando che i dati supportano un accoppiamento preattentivo dei canali somatosensoriale e uditivo nella rappresentazione corticale della frequenza.

Budd e Timora (2013) studiano il ruolo della corrispondenza temporale crossmodale di stimoli multisensoriali sull'integrazione corticale dell'informazione attraverso modalità sensoriali separate. Gli autori asseriscono che l'attività neurale oscillatoria nelle cortecce sensoriali potrebbe essere il meccanismo principale attraverso cui l'informazione sensoriale da modalità differenti viene integrata. Tramite le *steady-state EEG responses* (SSR) analizzano se le variazioni della corrispondenza temporale crossmodale dell'ampiezza modulata dalla stimolazione uditiva e tattile appaia nell'attività SSR della stimolazione multisensoriale (presentazione passiva e simultanea con modulazione uditiva e tattile). Gli autori scelgono la modulazione di 21 e 40 Hz per entrambe le modalità per massimizzare la risposta SSR. I risultati riscontrati corroborano le frequenze EEG corrispondenti alle stimolazioni uditive e vibro tattili e confermano come modulazione ottimale un'attività SSE diversa in base alla modalità, maggiore a 40Hz per risposte uditive e maggiore a 21Hz per risposte vibrotattili. Nonostante i cambiamenti consistenti e affidabili dell'attività SSR nelle manipolazioni della modulazione entro le modalità, i risultati non forniscono una forte evidenza delle interazioni multisensoriali nell'attività SSR per le condizioni crossmodali simultanee e non. Gli autori interpretano questi risultati come evidenza del possibile ruolo dell'attenzione come possibile fattore nel riconciliare i vari studi SSR sull'integrazione multisensoriale.

Sinestesia audio-tattile acquisita: il caso SR

La sinestesia è un fenomeno esperienziale di 'sovrapposizione' percettiva di due modalità sensoriali una volta che ne è stata attivata una (percepite distinte ma congiunte). Sebbene il legame tra sinestesia e corrispondenza cross-modale non sia ancora stato definito, per completezza riporto i dati presenti in letteratura in merito alla sinestesia audio-tattile, ovvero la percezione simultanea di un suono e una sensazione tattile. In letteratura non sono stati riportati casi di sinestesia audio-tattile pura e congenita, ma solo acquisita, ovvero conseguente a specifiche lesioni cerebrali.

Come riportato da J. Ward (2007) la maggior parte dei casi di sinestesia acquisita riguardano una deafferentazione sensoriale della modalità visiva, risultante in sinestesia audio-visiva

(Jacob et al, 1981 e Lessel e Cohen, 1979) e tattile-visiva (Armel e Ramachandran, 1999), in seguito a danno periferico delle vie visive (degenerazione della retina o danno del nervo ottico), ed un caso di sinestesia audio-visiva conseguente a tumore del mediale del cervello (REF). Questo caso presenta delle peculiarità notevoli (si veda la descrizione del caso riportata di seguito) e ha aperto la discussione sulle basi neurali e il legame tra le modalità tattili e uditiva. Nel 2007 Ro e coll. presentano per la prima volta un caso di sinestesia audio-tattile, in cui una paziente con lesione del nucleo ventrolaterale del talamo di destra (individuato tramite risonanza magnetica strutturale e DTI, *diffusion tensor imaging*) riporta sensazioni tattili in seguito a stimolazione uditiva. La paziente, successivamente chiamata 'caso SR', 36 anni, riporta questo tipo di sinestesia (una sensazione tattile durante un annuncio via radio) dopo 18 mesi dalla prima diagnosi neuropsicologica (neglect) conseguente ad un ictus cerebrale. Le sensazioni tattili riportate (formicolio o pressione) sono state costanti nel tempo (SR è stata nuovamente testata 6 anni dopo) ed esperite sempre nella parte superiore del corpo sinistra (mano e braccio). Gli autori riportano anche le differenze dell'esperienza di SR rispetto ai tipi di sinestesia non acquisita. In primo luogo, le esperienze sensoriali tattili riportati da SR sono molto semplici (formicolio e pressione), a differenza dei fenomeni sinestetici non acquisiti che sono solitamente esperienze sensoriali complesse (per esempio sinestetici che esperiscono forme piramidali indotte da sapori, o altri che esperiscono sapori ben precisi in seguito alla presentazione di parole). Questo è in linea con altre forme di sinestesia acquisita, che presentano anche fenomeni esperienziali semplici e dopo diversi mesi successivi al danno cerebrale. In secondo luogo, come accennato, non è stato trovato un tipo di sinestesia audio-tattile congenita (o dello sviluppo). Gli autori, inoltre, riportano i tre casi in cui sono state documentate esperienze tattili in seguito a stimolazioni sensoriali stimolando unaltra modalità sensoriale in sinestetici 'puri' (descritti da Luria, 1968, Cytowic, 2000, Simner et al. 2006, e and Simner e Holenstein, 2007).

Dal momento che in una fase iniziale (il primo anno e mezzo) SR ha riportato un fenomeno di 'antiestinzione' (in cui SR rileva eventi somatosensoriali controlesionali in modo significativamente più efficace, rispetto al lato ipsilesionale) gli autori (Ro et al, 2007) pensano che possa riflettere gli stadi iniziali dell'accoppiamento sensoriale e la riorganizzazione tramite il rinforzo delle connessioni callosali o l'alterazione delle connessioni talamiche. Aggiungono che la fase successiva in cui le sensazioni tattili illusorie tattili indotte dal suono possono riflettere vie talamo-corticali che si sono indebolite (visibile anche dai dati della DTI) o alterate come conseguenza della riorganizzazione della deafferentazione. Un'ipotesi che gli autori fanno, è quindi che la deprivazione degli input corticali del nucleo ventrale talamico abbia

rinforzato (o reso operative) di conseguenza alcune connessioni dirette tra le due cortecce uditiva e somatosensoriale, producendo il fenomeno sinestetico. Essi riportano, inoltre che visto che il percetto uditivo è presente anche in altre forme di sinestesia conseguenti a lesioni, è possibile riscontrare una spiegazione a livello neurale, nella fattispecie ipotizzare una 'cross-attivazione' (*cross-wiring*) nel cervello, di tipo subcorticale (Cytowic, 2000).

In generale, quindi, gli autori ipotizzano che la distruzione del nucleo ventrale talamico può quindi, aver causato cambiamenti nella connessione di regioni cerebrali remote, forse includendo un incremento di connessioni eccitatorie, o da connessioni intertalamiche alterate (o connessioni transcallocali che hanno origine nel talamo) che facilitano l'elaborazione dell'informazione sensoriale afferente quando il normale meccanismo di elaborazione è danneggiato. Ovviamente, il fatto che si tratti di una struttura come quella talamica, corrobora l'importanza e il coinvolgimento del talamo in diversi processi dell'elaborazione sensoriale in generale e il suo ruolo nella plasticità cerebrale (Ro et al., 2007).

In sintesi, quindi, gli autori riportano che queste connessioni alterate tra la regione talamica lesionata e la corteccia cerebrale, suggeriscono un ruolo di queste aree nel cambiamento della percezione sensoriale e una riorganizzazione della connettività assonale talamo-corticale che abbia contribuito al fenomeno. Interpretano quindi, questo dato in termini di una riorganizzazione sia funzionale che neurale delle regioni lesionate confinanti al nucleo ventrolaterale talamico.

Ward (2007) riprende il caso così come riportato da Ro e coll. (2007) e supporta l'ipotesi che si tratti di un dato a favore dell'esistenza di una plasticità compensatoria crossmodale, piuttosto che un 'smascheramento' di vie neuronali pre-esistenti, considerando sia la localizzazione del deficit, sia i casi precedenti di sinestesia acquisiti in cui si è iniziato a riportare il fenomeno. L'autore riporta inoltre, che il fatto che le sensazioni tattili riportate siano nella parte alta (piuttosto che bassa) del corpo, corrobora il principio di adiacenza', dato che la corteccia somatosensoriale che rappresenta il viso, le mani e le braccia è vicina alla corteccia uditiva. L'autore afferma che una terza alternativa possa riguardare l'esistenza di vie dirette audio-tattili che vengono rafforzate in caso di deficit neurologico o sensoriale.

Nel 2008 Beauchamp e Ro hanno nuovamente testato SR indagando con tecnica fMRI l'attivazione corticale alterata associata al danno talamico. Gli autori svolgono tre esperimenti: nel primo studio SR mostra risposta maggiore ai suoni nell'opercolo parietale (localizzazione della corteccia somatosensoriale secondaria-gli autori ipotizzano che questa anormale attività opercolare possa essere il substrato neurale correlato coinvolto in questo tipo di sinestesia-); nel secondo esperimento hanno verificato che i suoni che non evocano sensazioni tattili non

sono correlati con l'attivazione dell'opercolo, mentre i suoni che inducono la sensazione tattile riportano un'attivazione BOLD (potrebbe essere il risultato della plasticità indotta dalla perdita degli input somatosensoriali); nel terzo esperimento le risposte BOLD alla stimolazione somatosensoriale sono significativamente inferiori nei pazienti con opercolo piuttosto che nei controlli. Gli autori concludono che la doppia dissociazione della corteccia somatosensoriale secondaria nei pazienti (aumento delle risposte alla stimolazione uditiva e decremento delle risposte alla stimolazione somatosensoriale) suggeriscono che sia la plasticità indotta dall'ictus può risultare in una connessione anormale tra modalità sensoriali che sono normalmente separate , sia che la sinestesia può essere causata da connessioni inappropriate tra territori corticali vicini.

Nel 2008 Beauchamp e Ro svolgono tre esperimenti con tecniche MRI e fMRI, sempre con SR. Nei tre esperimenti viene analizzata l'attività neurale delle prove in cui vengono presentati degli stimoli visivi e uditivi (suoni naturali e artificiali) e solo stimoli visivi. Nel primo esperimento, gli autori hanno confrontato l'attività neurale di SR con nove controlli. Viene riscontrata attività neurale tre volte superiore, per SR rispetto ai controlli, dell'opercolo parietale (corteccia somatosensoriale secondaria). Nel secondo esperimento, quindi, gli autori si sono chiesti se l'opercolo parietale potesse essere il substrato neurale alla base della sinestesia di SR, ed effettivamente hanno verificato e riscontrato che i suoni che provocavano forti sensazioni sinestetiche tattili erano correlate con la maggiore attivazione dell'opercolo parietale, mentre quando venivano presentati suoni che non causavano la sinestesia non vi era alcuna attivazione dell'opercolo. Gli autori hanno ipotizzato che l'esperienza sinestetica fosse dovuta dalla plasticità conseguente alla perdita degli input somatosensoriali. Ragion per cui, nel terzo esperimento hanno riscontrato che SR mostrava attivazione inferiore dell'opercolo rispetto ai controlli in seguito a stimolazione somatosensoriale, come atteso. Gli autori concludono interpretando questi dati (doppia dissociazione delle risposte della corteccia secondaria somatosensoriale di SR, attivata da stimolazione uditiva e non somatosensoriale) a favore di una plasticità indotta dalla lesione che ha creato connessioni 'anormali' e 'inappropriate' di regioni corticali vicine e normalmente separate (Beauchamp e Ro, 2008).

Gli autori ipotizzano che cambiamenti di plasticità a breve e lungo termine tra la connettività delle cortecce uditiva e somatosensoriale hanno contribuito alla risposta opercolare uditiva di SR. Del resto, come accennato, la letteratura riporta di connessioni anatomiche tra le due modalità a stadi precoci dell'elaborazione corticale gerarchica (Schroeder et al., 2001).

Ipotizzano che un smascheramenteo a breve termine delle connessioni adiacenti uditive e tattili abbia contribuito alle risposte opercolari uditive, ma visto che il fenomeno si è presentato in

modo evidente dopo 18 mesi dal danno cerebrale, ipotizzano che cambiamenti di plasticità a lungo termine come la crescita assonale, possa avere giocato un ruolo più importante rispetto allo smascheramento (e quindi legato ai cambiamenti delle connessioni talamo-corticali). Gli autori ipotizzano che si possano essere verificati sia una crescita assonale, che un smascheramento di connessioni già esistenti, a favore di una plasticità neuronale. Aggiungono che questi cambiamenti hanno causato l'attivazione compensatoria funzionale portando alla sinestesia di SR (Rouw and Scholte, 2007). Suggeriscono inoltre, che se SR nel periodo successivo alla lesione, avesse svolto della riabilitazione mirata alla stimolazione somatosensoriale, probabilmente avrebbe prevenuto la formazione di connessioni inappropriate e non avrebbe sviluppato questa forma di sinestesia.

Naumer e van den Bosch (2009) condividono l'interpretazione di Beauchamp e Ro, (2008) e sottolineano che le proiezioni talamo corticali possano avere un ruolo importante per prevenire la percezione cross modale illusoria e per sottolineare l'integrazione multisensoriale affidabile.

Nel 2013 Ro e coll. Hanno svolto un altro studio con la paziente SR con tecnica MRI. Dal momento che gli studi precedenti avevano mostrato che la corteccia somatosensoriale di SR diventava più responsiva gli stimoli uditivi, visto che non riceveva gli stimoli somatosensoriali (Ro et al. 2007; Beauchamp and Ro 2008), gli autori hanno ipotizzato che ciò potesse essere dovuto ad un rafforzamento di normali connessioni audio-somatosensoriali, che si siano rinforzate a cause della mancanza dei normali input talamici. Ragion per cui, in questo studio hanno comparato la densità delle connessioni della corteccia uditiva e le diverse sottodivisioni della corteccia somatosensoriale primaria e secondaria, tra SR e dei controlli.

Gli autori hanno riscontrato l'esistenza di molte connessioni ipsilaterali tra le cortecce primaria uditiva e somatosensoriale primaria e secondaria in tutti i soggetti. La paziente SR, però, mostra molte più connessioni tra la corteccia uditiva e la corteccia somatosensoriale secondaria nell'emisfero lesionato, base neurale collegata con l'elaborazione percettiva dei suoni (che elicitava sensazioni corporee). Gli autori concludono dicendo che queste potrebbero essere le basi anatomiche tra l'interazione audio-tattile e che entrambe le modalità possono influenzarsi vicendevolmente. In sintesi, gli autori hanno ipotizzato che in seguito alla lesione talamica destra, la corteccia somatosensoriale destra non ricevesse più input somatosensoriali diretti, e questo è stato dimostrato dalla sua mancata sensibilità e dallo scan DTI (Ro et al. 2007). In seguito a ciò, le connessioni tra A1 e S1 sono diventate più rade nel lato lesionato, ma le connessioni tra A1 e OP1/S2 sono diventate più reattive alle connessioni audio-somatosensoriali già esistenti. Concludono ipotizzando che i danni cerebrali, direttamente

connessi alle regioni della lesione può degenerare, ma anche regioni connesse indirettamente possono essere coinvolte nella riorganizzazione per compensare il danno.
Ipotizzano perfino che vi possa essere una origine comune tra le due modalità: la prossimità e connettività tra A1 e la corteccia somatosensoriale secondaria, (si veda l'ipotesi di una simile elaborazione di frequenze di base che presentano meccanismi comuni con comuni origine genetiche Ladher et al. 2010), fa supporre, sostengono che il senso dell'udito sia evoluto da quello del tatto, permettendo l'estrazione di informazioni dello spazio extrapersonale.

Per completezza, riporto un altro caso 'simil' sinestetico audio-tattile. Nel 2011 Fornazzari e coll. hanno riportato un altro caso di un paziente di 45 anni, iperteso, che dopo 9 mesi dopo un danno emorragico del nucleo laterale posteriore talamico sinistro, ha manifestato vari tipi di sinestesia, tra cui la sinestesia audio-tattile. Il paziente ha riportato che perfino pensare ad una stimolazione sensoriale può causargli l'esperienza sinestetica. Da notare, che a differenza di SR, questo paziente presenta il 'trigger' e 'concurrent' invertiti, ovvero l'ascolto di suoni alti (come musica Hindi o cinese) provocano sensazioni extracorporee quali ' cavalcare, fare surf o volare sulla musica'. Il fatto che il paziente riporti di poter volontariamente sopprimere l'esperienza sinestetica e che in alcune situazioni non esperisca il fenomeno, mi porta ad escludere il caso come buon esempio di sinestesia, visto che non rispetta i requisiti di sistematicità e irrefrenabilità/incontrollabilità del fenomeno sinestetico (Cytowic, 2002).
Ad oggi, non sono presenti altri casi di sinestesia audio-tattile. Il fatto che essa risulti solo come conseguenza di un danno cerebrale, porterebbe a pensare al fenomeno come frutto di una plasticità cerebrale. Il fatto che riveli connessioni uditive e somatosensoriali, come abbiamo visto, non è un dato sorprendente in natura, ma il fatto che non vi siano casi puri potrebbe indicare un legame 'non immediatamente comune' tra le due modalità sensoriali. Le evidenze sperimentali sulla sinestesia audio-tattile corroborano gli studi di neuroimmagine riportati in precedenza e la possibilità di interazione tra queste due modalità. Tuttavia, come si vedrà nella conclusione e dai risultati di questo lavoro, probabilmente il legame tra queste due modalità richiede una specializzazione neuronale ben specifica, conseguente a deprivazione sensoriale in seguito a danno cerebrale, o associazioni apprese durante l'arco di vita.

1.3 Sostituzione sensoriale

Il termine 'sostituzione sensoriale' indica letteralmente la sostituzione, in termini funzionali, di una modalità sensoriale con un'altra. Per sostituzione 'in termini funzionali' si indica che un'altra modalità rispetto a quella usuale, trasmette lo stesso tipo di informazioni, che possano determinare una risposta comportamentale altrettanto efficace di interazione con l'ambiente. In altre parole, la teoria della sostituzione sensoriale si basa sull'idea che l'equivalente funzionale di una modalità sensoriale possa essere ottenuto in assenza di quell input sensoriale specifico, se quella informazione viene convertita in un codice di un'altra modalità sensoriale, in un modo che riesca a preservare le caratteristiche fondamentali del segnale originale (rivisitazione di Proulx, Brown, Pasqualotto e Meijer, 2014).

Occasionalmente, il termine può essere considerato come un'espressione racchiudibile nel concetto di plasticità neurale, che la include. Essa infatti, può essere riferita alla flessibilità in cui una capacità sensoriale intrinseca viene usata e legata ad un'altra capacità sensoriale (Visell, 2008). In questa prospettiva, la sostituzione sensoriale offre la possibilità di recupero di abilità (per esempio in popolazioni con deficit sensoriali) attraverso la plasticità neurale che consente una 'risposta adattiva ad una domanda funzionale' (Bach-y-Rita e coll., 2003).

Se una modalità sensoriale deve veicolare lo stesso tipo di informazioni, o meglio informazioni equivalenti funzionalmente, allora bisogna considerare che la sostituzione sensoriale implica un rimappaggio dell'informazione sensoriale in entrata (stimolo di input), la sua interpretazione a livello percettivo e la conseguente risposta motoria. Jóhannesson e coll. (2016) per dare un'idea di questo concetto e più in generale della capacità di flessibilità del cervello nel ricalibrarsi continuamente, portano l'esempio dei touchpad della tastiera del computer. Il cervello degli utenti deve costantemente svolgere due conversioni: una che riguarda il movimento del dito con il movimento visivo, e l'altra che riguarda il piano, ovvero convertire

un movimento che avviene in una superficie orizzontale, nel movimento che visualizzano sullo schermo in verticale.
Un altro esempio di un'esperienza di sostituzione sensoriale comune è quando ci troviamo in un ambiente caotico e un nostro amico prova a parlare con noi: in questa situazione la lettura del suo labiale ci darà l'impressione di 'sentire meglio' cosa sta dicendo.
Ovviamente, questa 'sostituzione dei sensi' è strettamente legata ad una stimolazione esterna, legata per lo più ad un apprendimento del cervello ad usare un'area al posto di un'altra. L'apprendimento ad usare un'altra modalità sensoriale viene per lo più reso possibile dall'utilizzo di sistemi di sostituzione sensoriale o *sensory substitution devices* (SSDs), ovvero delle interfacce uomo-macchina non invasive, che usano algoritmi che transcodificano informazioni sensoriali di un tipo in un altro tipo. I sistemi di sostituzione sensoriale o *Sensory substitution devices* (SSDs), presuppongono che il cervello sia capace di utilizzare un'informazione proveniente da un'altra modalità sensoriale e che sia capace di ricalibrarsi in tempo reale, in modo tale che tale informazione possa essere utilizzata come se si trattasse di un'altra modalità sensoriale (Jóhannesson e coll., 2016). Dopo un buon periodo di training, un dispositivo di sostituzione sensoriale riuscirà a trasmettere un'esperienza percettiva 'genuina' attraverso un 'nuovo' senso. Un'esperienza continua porterà quindi ad una risposta automatica simile a quella esperita normalmente attraverso un'altra modalità sensoriale, esperita in modo naturale. Ciò, come già menzionato, si riallaccia sia agli studi sulla plasticità cerebrale che sugli studi di adattamento (training a breve e a lungo termine) che il cervello è in grado di fare per ottenere una flessibilità funzionale efficace.

Per inciso, sistemi di 'conversione sensoriale' sono stati utilizzati, per la verità, già in epoca napoleonica, per poter leggere al buio negli accampamenti; ideato da Barbier e poi sistematizzato da Louis Braille, che codificò il sistema di lettura ormai ben noto, utilizzato da non vedenti. Successivamente, negli anni '50 sono stati utilizzati mezzi di conversione di testi più sofisticati, come l'Optacon (Goldish and Taylor, 1974), ma per far si che l'utilizzo degli sistemi di sostituzione sensoriale possa supportare la teoria della sostituzione sensoriale, occorre che essi, oltre a produrre una risposta funzionale, siano accompagnati da una genuina esperienza di percepire un'esperienza sensoriale tramite un'altra (si veda il paragrafo successivo).

Il problema della sostituzione sensoriale, quindi, sembra essere fortemente collegato con quello dell'apprendimento, come disquisiscono Proux e coll. In una review del 2014, se la plasticità neurale e i cambiamenti neuroanatomici sono legati all'apprendimento di nuovi sistemi di sostituzione sensoriale, allora la chiave fondamentale diventa capire come

l'apprendimento possa essere generale o specifico (si veda paragrafo sulle critiche alla sostituzione sensoriale).
L'importanza dello studio di questa teoria presenta non solo interesse circa la percezione umana e l'apprendimento percettivo, ma ha notevole importanza da un punto di vista applicativo, per popolazioni con deficit e deprivazione sensoriale.

Cosa ci rende ottimisti rispetto alla sostituzione sensoriale? Evidenze sperimentali a favore.

Vi sono oramai parecchi studi a supporto di questa teoria, sia a livello comportamentale che neuronale.
I primi tentativi sono stati svolti con l'intento di aiutare popolazioni quali i non vedenti, per cui si è cercato di permettere una 'visione' funzionale in assenza di input visivo, ovvero si è cercato di veicolare informazioni visive tramite una rappresentazione che potesse essere codificata tramite stimoli sonori o tattili dando la possibilità di 'vedere' attraverso le orecchie e la lingua (Bach-y-Rita et al., 1969; Meijer, 1992).
Durante uno studio pionieristico del 1969, Bach-y-Rita e coll. (risultati replicati nello studio del 1972) hanno sviluppato uno strumento che dava feedback aptici per trasferire informazioni circa l'ambiente circostante. Tramite le vibrazioni di una sedia (trasmesse alla schiena dei partecipanti), sia partecipanti non vedenti e vedenti bendati, dopo un periodo di training attivo, sono riusciti a codificare l'ambiente circostante. Ciò che rende rilevante questo studio è il fatto che i partecipanti hanno riferito di percepire/sentire le informazioni come provenienti dall'ambiente, piuttosto che dalla sedia.
Uno dei sistemi di sostituzione sensoriale o *sensory substitution device* (SSD) che ha avuto maggiore diffusione è lo strumento chiamato "The vOICe", letteralmente OIC "Oh, I see" (Meijer, P.B.L., 1992). Nella sua versione portatile (Meijer, 1998), il vOICe si presenta come un comune occhiale a cui è attaccata una telecamera, ed è disponibile anche un'applicazione (chiamata *EyeMusic*, Amedi e coll., 2007) scaricabile per l'iPhone. Lo strumento transcodifica informazioni visive in suoni. In particolare, nella sua versione più recente, vengono associati suoni di diversa altezza ad altezze differenti (trasmettendo così la forma dell'oggetto) e il volume con la luminosità (colore dell'oggetto) in uno scanning continuo da sinistra a destra. Dopo diverse settimane di training i soggetti hanno riportato di avere la sensazione di 'vedere' tramite i suoni. Inizialmente usato per orientarsi in ambienti, anche piccoli, come stanze, adesso

viene utilizzato anche per il riconoscimento di volti e posizioni di persone in fotografie (Amedi e coll., 2007). Una cieca acquisita che ha avuto la vista per circa vent'anni, ha riportato in un'intervista che nell'arco di due-tre settimane ha sviluppato un senso della 'forma del suono' e dopo tre mesi ha avuto l'impressione di vedere dei flash dell'ambiente circostante e di riuscire a identificare oggetti guardando a loro, "come si vede" (Pat Fletcher, in ACB Braille forum). Anche a livello neuronale sono stati riportati dati significativi in seguito a training di persone non vedenti (vedi sotto).

S. Maidenbaum e coll. (2014) riportano nella loro diversi compiti portati a termine con successo in seguito a training (per lo più riguardano sostituzione sensoriale della modalità visiva con stimolazione uditiva o tattile): evitamento di ostacoli, percezione della profondità, detezione di movimento, raggiungimento di oggetti, riconoscimento di oggetti, stima della dimensione relativa, identificazione della posizione degli oggetti, acuità visiva funzionale.

Altri risultati a favore vengono riportati da Amedi et al., 2007; Auvray et al., 2007; Matteau et al., 2010; Renier and De Volder, 2010; Stiles et al., 2012; Wright et al., 2012.

A livello neuronale, si è riscontrato che aree corticali primarie normalmente collegate all'attivazione di alcune modalità sensoriali, vengono attivate anche da altre modalità sensoriali (per esempio, Maidenbaum e coll., 2014).

Innanzitutto, bisogna tenere presente che molti studi su popolazioni con deprivazione sensoriale riportano attivazione di aree sensoriali comuni anche alle popolazioni senza deficit sensoriale. Un esempio è lo studio di Sadato e coll. (REF) hanno mostrato in uno studio con tomografia ad emissione di positroni (PET) che la lettura Braille attiva la corteccia visiva secondaria, nella popolazione non vedente. Hanno quindi chiesto ai partecipanti di svolgere compiti tattili non-Braille e confrontare le aree attivate con quelle attivate durante la lettura Braille. I partecipanti hanno mostrato attivazione delle aree visive primarie nei compiti di discriminazione tattile, mentre partecipanti congeniti hanno mostrato attivazione opposta. Questi risultati, ma vi sono anche molti altri studi in merito, hanno mostrato attivazione della corteccia visiva in non vedenti per compiti tattili.

Le evidenze in merito alle popolazioni speciali hanno supportato la creazione di strumenti di sostituzione sensoriale che potessero mostrare la plasticità corticale in popolazione con deprivazione sensoriale, anche in seguito a training di utilizzo di questi strumenti. Alcune evidenze sono state riportate in studi con i ciechi, i quali, dopo un periodo di training hanno mostrato attivazione di aree occipitali. Un esempio in merito riguarda l'adattamento a livello cerebrale è stato riscontrato dal gruppo di ricerca che ha sviluppato il vOICe. Amedi e coll. (2001) hanno misurato fisiologicamente i segnali fMRI di un gruppo di ciechi acquisiti che ha

mostrato una notevole attivazione delle aree deputate all'analisi visiva e una stimolazione di queste aree che impediva l'attività della corteccia visiva, ostacolava anche l'esplorazione quando usavano il vOICE (Merabet et al., 2009)

Critiche e limiti della teoria della sostituzione sensoriale.

Una delle critiche più forti contro queste evidenze a favore della sostituzione sensoriale, riguarda la supposizione che questi 'cambiamenti di percezione' siano legati al tipo di compito appreso e pertanto non sia corretto parlare di sostituzione sensoriale in toto, ma semplicemente 'associazione di nuove risposte' per un compito. In altre parole, se gli utenti non sono in grado di generalizzare questo 'sentire attraverso un altro senso', allora questo presenta un limite per poter parlare di sostituzione sensoriale. In sintesi, quindi, devono tenersi in considerazione fattori importanti quali il rilevamento attivo e la chiusura del loop senso-motorio (Reynolds e Glenney, 2012) o l'acquisizione di abilità di generalizzazione (Kim e Zatorre, 2008).

Inoltre, l'assenza di programmi di training ben stabiliti, influenza negativamente l'impatto e l'utilizzabilità dei sistemi di sostituzione sensoriale, portando anche a dati contrastanti e meno convincenti. Ciò risulta essere rilevante se si considera il grande sforzo richiesto agli utenti prima che si verifichi un reale vantaggio, o meglio, una sensazione di 'esperienza percettiva' reale, legata ai tempi di adattamento con lo strumento (nonché i tempi di adattamento cerebrale o di plasticità cerebrale).

Fatte queste premesse, di seguito verrà riportata una teoria di sostituzione sensoriale alla quale si è fatto riferimento nella parte sperimentale del presente lavoro e verranno riportati in dettaglio, gli studi sui sistemi di sostituzione sensoriale tattili, ovvero i tentativi che sono stati fatti fino oggi per trasmettere informazione linguistica tramite strumenti tattili.

1.3.3 *Sensory motor theory.*

Dal 2001 Kevin O'Regan ha proposto un nuovo modo di approcciarsi al come e al perché noi 'sentiamo' il mondo nel modo in cui lo sentiamo. Domandandosi il perché una modalità sensoriale dà una specifica sensazione piuttosto che un'altra l'autore (prima riferendosi alla visione e poi estendendo la questione anche alle altre modalità sensoriali) suggerisce un modo alternativo alla fenomenologia di come le modalità sensoriali funzionano, formulando la teoria motoria sensoriale, o *Sensory Motor Theory*. L'approccio sensomotorio, dichiara che ciò che la gente intende dicendo che qualcosa ha un 'sentire' è strettamente legato all'interazione motoria. L'autore indica quattro proprietà distintive delle interazioni sensoriali. 1) Innanzitutto,

le interazioni sensoriali sono ricche, nel senso che possono assumere una molteplice varietà di stati possibili. 2) Inoltre, a differenza degli altri processi cerebrali, gli input sensoriali variano sistematicamente a seconda di come si muove il corpo (*bodiliness*), 3) e talvolta possono cambiare di per sé (parziale *insubordinateness*). 4) Le interazioni sensoriali inoltre, in alcune circostanze possono catturare l'attenzione (per esempio, quando gli input sensoriali cambiano improvvisamente).

L'autore afferma che l'approccio sensomotorio fornisce le descrizioni base dell'esperienza qualitativa delle sensazioni specifiche che sono accessibili al pensiero e al linguaggio ordinario. Di contro, afferma che i canali neurali devono essere descritti in termini di attività neurale, per la quale non c'è un legame naturale con le descrizioni usate ogni giorno per descrivere l'esperienza (O'Regan, 2011).

In quest'ottica, gli strumenti di sostituzione sensoriale mostrano come sia possibile stimolare una modalità sensoriale e 'sentire qualcosa' in un'altra modalità, il che conferma che ciò che conta nel determinare la natura dell'esperienza di una sensazione non è la parte cerebrale che viene direttamente stimolata, ma la modalità di interazione che è implicata nell'interazione con l'ambiente. L'autore riporta che se durante il loro sviluppo i furetti usano la corteccia uditiva per processare informazioni visive, imparano a interagire con l'ambiente in modo funzionale usando questo tipo di informazione (avranno quindi, una sensazione visiva nonostante utilizzino la corteccia uditiva). Come si è visto, in modo similare anche i ciechi congeniti che in seguito a stimolazione uditiva presentano attività nella corteccia occipitale (normalmente deputata all'elaborazione visiva) (Amedi et al, 2007). Egli conclude quindi, che il ruolo del cervello in tutto questo è quello di servire come substrato che impara modelli appropriati di interazione sensomotoria. Aree corticali diverse potrebbero essere più o meno specializzate in questo compito e apprendere in modo più o meno efficiente. Ma nel momento in cui l'apprendimento è fatto, la sensazione corrisponderà alla modalità di interazione coinvolta.

Progetto "*Feelspeech*": ascoltare tramite la pelle.

Nella parte sperimentale del presente lavoro, si farà riferimento all'analisi e utilizzo di uno strumento di sostituzione sensoriale che è stato sviluppato dal team del prof. Kevin O'Regan (*Laboratoire Psychologie de la Perception*) all'interno del progetto convenzionato dall'Unione Europea, ERC (European Research Council) denominato *FeelSpeech project.*

Il progetto *FeelSpeech* è nato come implicazione pratica della teoria sensomotoria (si veda il paragrafo precedente per analogie e differenze con l'approccio di sostituzione sensoriale più generale). Il progetto *FeelSpeech* nasce dalle potenzialità teoriche implicite di questa teoria,

supponendo che una persona con problemi d'udito potrebbe avere l'impressione di sentire parlare (e quindi un decremento del deficit uditivo) tramite l'utilizzo di uno strumento chiamato *FeelSpeech*, letteramente 'ascoltare tramite il sentire' (inteso come sensazione tattile). In pratica, a livello percettivo ci si aspetta un incremento della comprensione del linguaggio, in assenza di un incremento dell'informazione uditiva, ma tramite l'utilizzo dell'informazione tattile che viene in qualche modo 'inglobata' o semplicemente 'veicola' informazione simile, o funzionalmente pertinente, a quella tattile. L'ipotesi fondamentale alla base del progetto è che ci sia un'integrazione automatica tra la stimolazione tattile e l'input uditivo di scarsa qualità e dopo un ripetuto utilizzo del dispositivo, una persona con perdita parziale dell'udito, 'sente' meglio il linguaggio, o in altre parole, integra nella percezione del parlato l'informazione tattile ed esperisce una sensazione di ascoltare meglio. In questa prospettiva, dopo molta pratica, il dispositivo di sostituzione sensoriale (in questo caso, come già detto, il cosiddetto *Feelspeech*) trasmetterà un'autentica esperienza "percettiva" del "nuovo" senso. Lo scopo di questo progetto è quello di aiutare (facilitare o disambiguare) le persone con problemi all'udito al fine di comprendere meglio il linguaggio.

Il progetto FeelSpeech ha, quindi, lo scopo di aiutare le persone con problemi all'udito (che avevano un'esperienza di audizione normale precedente) per "ascoltare" meglio le informazioni del discorso attraverso un dispositivo tattile chiamato FeelSpeech.

Ovviamente, lo sviluppo di tale strumento non è l'oggetto principale del presente lavoro, che si presenta come un contributo di ricerca di un'applicazione dello stesso.

Perché usare il tatto per veicolare informazione uditiva e in particolare quella linguistica?

Come affermato da Kirman (1973), il senso del tatto è stato considerato per lungo tempo un cattivo candidato per trasmettere le informazioni del linguaggio, per diversi motivi: 1) tutti i tentativi precedenti di costruire con successo un dispositivo tattile per la comprensione del linguaggio sono stati deludenti; 2) il linguaggio è un codice speciale, ovvero univocamente e biologicamente legato al sistema uditivo, per cui la pelle (così come ogni altra modalità sensoriale) è soggetta a fallire nella sua decodifica; infine, 3) il potere risolutivo della pelle, sia temporalmente che spazialmente, è troppo limitato per affrontare le complessità del linguaggio. Per quanto riguarda la prima affermazione, come vedremo, probabilmente la maggior parte dei precedenti dispositivi esistenti è stata abbastanza scarsa sia in termini di estrazione del segnale vocale (in realtà ancora non sappiamo esattamente quali tipi di informazioni usiamo per decodificare il linguaggio), sia per il tipo di stimolazione che essi utilizzavano (vibrazione e leggere scosse elettriche hanno un effetto di desensibilizzazione della pelle e non sono affatto

piacevoli). Per quanto riguarda il secondo punto, oggi esistono sistemi abbastanza buoni che possono trasmettere on line il linguaggio utilizzando altri canali sensoriali (per lo più è stato utilizzato il canale visivo, ad esempio occhiali o applicazioni google o il linguaggio cosiddetto 'Cued', inventato nel 1966 dal dottor R. Orin Cornett). Per quanto riguarda l'ultima affermazione, il potere della pelle in termini di risoluzione temporale e spaziale è stato ampiamente sottostimato per lungo tempo (per studi sistematici temporali vedere Gault, 1924, 1926a, 1926b, 1927).

La difficoltà indicata al primo punto è stata in parte superata dall'utilizzo di un tipo di stimolazione che abbiamo denominato *microswiping*, ovvero l'utilizzo di una stimolazione completamente differente rispetto alla vibrazione o all'elettricità, in quanto prevede la sensazione di 'una linea' in movimento (si veda la descrizione completa dello strumento nella sezione sperimentale). In questo modo, la desensibilizzazione della pelle viene posticipata di gran lunga e non risulta spiacevole.

In risposta al secondo punto si può notare che finora i sistemi efficaci nel veicolare informazioni linguistiche sono stati di tipo visivo, nonché il canale sensoriale più ampiamente studiato in letteratura (erroneamente pensato come il canale sensoriale primario dell'uomo). Di fatto, la maggior parte delle persone con deficit uditivo, riporta stanchezza agli occhi, costantemente concentrati per 'interpretare' il labiale delle persone che le circondano, ragion per cui un effettivo aiuto nella comprensione del linguaggio utilizzando altre modalità sensoriali, quali il tatto, risolverebbe questo problema. Il fatto che l'informazione linguistica è stata codificata tramite questi sistemi, non implica un'effettiva sostituzione sensoriale, per la quale ancora si sono svolti pochi studi (in nessuno dei sistemi descritti sopra si ha una percezione di 'incremento' della comprensione linguistica, ma i soggetti si limitano a leggere le informazioni linguistiche captate dai sistemi e tradotte).

Riguardo al terzo punto, si ricordino gli studi svolti da Geldard (1957, 1960), in cui il tatto è stato studiato sistematicamente come mezzo per la comunicazione. Egli ha messo in evidenza le diverse possibilità di tocco: (a) la capacità della pelle nelle discriminazioni temporali e spaziali, simile a quelle dei nostri occhi e orecchie; b) l'efficacia delle sensazioni cutanee nel catturare l'attenzione; (c) l'area potenziale stimolabile della pelle (circa 1,8-2,0 m² nell'area dell'adulto medio; Montagu, 1971); e (d) la scarsa utilizzazione di questo canale per la presentazione delle informazioni. Non per nulla, Geldard (1957) ideò il primo linguaggio tattile, il *Vibratese*, basato sia sulle considerazioni pratiche che sui risultati di una serie di esperimenti psicofisici controllati sulla discriminazione tattile (cfr. paragrafo "Dispositivi tattili per la comprensione del linguaggio"). Gli studi di Geldard e la sua applicazione pratica, hanno

dato speranza per un effettivo impiego della modalità tattile nel veicolare informazione linguistica.

Comprendere il linguaggio tramite le mani: il metodo di Tadoma.

Una delle prime applicazioni del tatto per la comprensione linguistica è stato il metodo di Tadoma negli anni ‘20. Il Tadoma è un metodo che è stato utilizzato (ed evoluto) inizialmente all'interno di comunità di sordomuti (Reed et al., 1985). I sordomuti mettevano una mano sul volto della persona che parla. Tipicamente, il pollice poggiava leggermente sulle labbra di chi parla, mentre le altre dita poggiavano sulla guancia e sulla mascella. Questa posizione consentiva di monitorare diverse azioni facciali relative alla produzione vocale, quali movimento delle mandibole, flusso d'aria e vibrazioni delle corde vocali. I sordomuti che hanno fatto ampio uso di questo metodo, riuscivano a comprendere il discorso in modo del tutto similare alla velocità di un 'ascolto normale. L'uso del Tadoma è stato più frequente tra il 1930 e il 1960 in Nord America, dove è stato insegnato nelle scuole per i bambini sordi (Reed, 1996). Alcuni sforzi sono stati fatti per sviluppare versioni semplificate del Tadoma per la presentazione di segnali registrati (che simulano i movimenti del volto di un parlante). Reed et al. (1985) ha introdotto un volto artificiale che imitava le azioni facciali della vibrazione, del flusso d'aria orale, del movimento delle mascelle e del movimento delle labbra. Uno studio sulla presentazione di singole lettere ha indicato che il Tadoma semplificato potrebbe fornire una simulazione grossolana del Tadoma naturale (Leotta, Rabinowitz, Reed e Durlach, 1988). Oggi il metodo di Tadoma è ancora abbastanza utilizzato per la popolazione cieca e sorda, ma è meno diffuso del Braille (Pasquero, 2006). Il sistema risulta essere, come facilmente intuibile, piuttosto ‘intrusivo’ e informale, il che lo rende poco adatto e flessibile ai contesti formali.

Dispositivi tattili per la comprensione del linguaggio: background storico

L'idea di utilizzare il senso del tatto per comprendere il linguaggio si è sviluppato sui risultati delle ricerche sulle tecnologie di visualizzazione tattile, che sono sempre più sofisticate, meno intrusive e quindi più efficaci e accettabili per gli utenti.

Si tenga presente che, come distinguerebbe Sherrick (1984), a differenza del metodo Tadoma che riflette un linguaggio ‘naturale’ (ovvero un linguaggio che non richiede una ricodifica per poter essere percepito dall’ascoltatore), i dispositivi tattili utilizzano un linguaggio ‘sintetico’ (ovvero implicano che il messaggio vada ricodificato).

Di seguito verranno descritti i dispositivi tattili precedentemente usati per trasmettere le informazioni linguistiche, focalizzandoci sul perché non hanno avuto successo e sono stati abbandonati.

La letteratura sui dispositivi tattili (dispositivi vibratori o elettrici) per aiutare nella comprensione del linguaggio è stata particolarmente intensa agli inizi del XX secolo.

La prima versione di un dispositivo tattile fu fatta nel 1924 da Gault, il quale ha usato un tubo lungo 14 piedi (circa 4,27 m). In una estremità, il parlante poggiava la bocca nel tubo che non faceva altro che condurre le vibrazioni della voce di chi parlava sul palmo della mano dell'ascoltatore che teneva la mano nell'altra estremità del tubo. L'isolamento acustico assicurava che il soggetto non potesse udire alcun suono. Dopo parecchia pratica, Gault riporta che l'ascoltatore poteva distinguere tra di loro, diverse parole.

Nel 1927 Gault sviluppò il Teletactor, un sistema di sostituzione sensoriale che usava vibrazioni per facilitare i sordi nella comprensione del discorso. Il sistema era basato su un'idea di conversione diretta di diverse frequenze del linguaggio umano in vibrazioni tattili. Il sistema aveva cinque vibratori collegati alle dita dell'utente. Ogni vibratore presentava una diversa banda di frequenza del discorso. Ad esempio, la lettera "a" con frequenza ascendente è stata prima sentita sul pollice e poi sul dito anulare. Gli studi condotti da Goodfellow (1934) hanno mostrato che il Teletactor riusciva a migliorare la comprensione della lettura del linguaggio labiale dei bambini sordi che usavano il dispositivo in concomitanza del linguaggio labiale. Di contro, studi successivi hanno dimostrato che solo poche persone erano effettivamente in grado di interpretare il linguaggio usando la sola vibrazione. Come notato da Geldard (1957, 1960), la limitazione principale del sistema è stata causata dal fatto che non teneva conto delle proprietà della pelle umana. Egli sostiene che le frequenze implicate nel linguaggio umano sono tipicamente tra 200 e 3500 Hz (Kirman, 1973), mentre la gamma di frequenza del senso cutaneo varia tra 0,3 e 500 Hz (Goldstein, 1999). Pertanto, vi era una mancata corrispondenza tra la stimolazione presentata e le capacità della gente di percepirla.

Durante il 1949 e il 1950, Wiener et al. hanno sperimentato un guanto (Felix) collegato a un *vocoder* (un dispositivo utilizzato per trasmettere messaggi su cavi telefonici a lunga distanza, in altre parole, è il sintetizzatore che produce suoni da un'analisi dell'input vocale). È stato usato solo poche volte prima che i ricercatori lo abbandonassero.

Nel 1957, Gerald, come già citato, ha ideato il *Vibratese.* Vibratese, è nato in seguito alle scoperte degli esperimenti psicofisici sulla percezione tattile. A differenza degli altri linguaggi tattili, invece di utilizzare la frequenza di vibrazione come parametro, utilizzava la variazione dell'ampiezza, della durata e della posizione del corpo di stimolazione. I tre diversi livelli di

ampiezza e tre durate sembravano essere riconosciuti molto bene dagli utenti. Cinque attuatori al petto dell'utente consentivano di variare la posizione spaziale degli stimoli. Da questi tre parametri ne sono conseguiti 45 diversi tipi di stimoli, che erano sufficienti per tutte le lettere, numeri e alcune parole brevi. Per ottimizzare la velocità di presentazione e per evitare confusione, le lettere più frequenti sono state assegnate alle durate più brevi e alle diverse posizioni spaziali. Gli esperimenti che utilizzavano la vibrazione hanno mostrato che erano necessarie molte ore di formazione per raggiungere livelli di lettura soddisfacenti. Secondo Geldard, gli utenti avevano bisogno di circa 12 ore per imparare l'associazione segnale-simbolo per poter leggere le parole. In uno studio successivo, Geldard (1960) constatò che per arrivare a capire le frasi composte da parole a cinque lettere (presentate ad una velocità di 38 parole al minuto) con un'accuratezza del 90%, gli utenti avevano bisogno di 65 ore di apprendimento. Ipotizzò che una velocità teorica di 67 parole al minuto potesse essere raggiunta diminuendo il tempo necessario per presentare le lettere. Tuttavia, questo in pratica non poteva essere effettuato, dal momento che l'incremento della velocità di presentazione avrebbe richiesto un codificatore automatico invece di una presentazione manuale (come disponibile all'epoca). Anche se il Vibratese è stato uno dei sistemi di comunicazione primordiali di maggior successo, non ha mai guadagnato forte popolarità ed è stato completamente abbandonato.

I successivi tentativi di costruire strumenti che usavano le vibrazioni come tipo di stimolazione, sono stati abbandonati. Si ricordano tra gli altri, i sistemi ideati da Guelke e Huyssen (1959), Kringlebotn (1968) e Pickett e Pickett (1963) i quali hanno usato sofisticati vocoders tattili, ma hanno osservato che le prestazioni dei loro soggetti non erano incoraggianti e hanno abbandonato le ricerche.

Uno dei fattori per cui questi dispositivi non sono risultati efficaci, potrebbe essere legato al fatto che non erano in grado di fornire le informazioni dettagliate necessarie per un'adeguata percezione del linguaggio. In quest'ottica è possibile che vi fosse un limite sia nel tipo di informazione fornita, che nel tipo di stimolazione presentata (difficile da comprendere o limitata temporalmente la disponibilità in seguito alla desensibilizzazione della pelle)

Un altro fattore potrebbe essere correlato al tipo, e quindi, all'inadeguatezza della pratica o training svolto antecedentemente al loro utilizzo. Il gruppo di ricerca di Clark ha enfatizzato l'importanza del programma del training e la sua importanza sull'efficacia degli ausili tattili per la comprensione del parlato (vedi Galvin et al. 1993) portando qualche evidenza in merito (si veda il paragrafo sul training).

Dispositivi tattili per la comprensione del linguaggio: caratteristiche tecniche.

Come accennato, i vari dispositivi tattili creati fino ad oggi, differiscono per 1) tipo di informazione linguistica (frequenze) che i vari autori hanno deciso di convertire in input tattile (e di conseguenza, il numero di canali necessari per veicolarle), 2) area in cui viene effettuata la stimolazione, 3) tipologia di stimolazione.

Riguardo alla tipologia di informazione linguistica estrapolata e veicolata, in letteratura i diversi autori hanno scelto di veicolare diverse frequenze e si distinguono i dispositivi a uno-due canali da dispositivi multicanale. In base al tipo di utenza e agli obiettivi prefissati, negli anni, si è visto che l'utilizzo di un solo tipo di frequenza riusciva ad essere molto efficace nel veicolare alcuni tipi di informazioni linguistiche che 'davano' l'impressione di facilitare la comprensione linguistica, ma non era sufficiente per veicolare tutte le informazioni linguistiche, per esempio aiutare i sordi profondi nella comprensione linguistica.

I dispositivi a uno o due canali sono abbastanza semplici. Presentano come vantaggio quello di aiutare le persone con udito parzialmente preservato, a seguire il ritmo del linguaggio, e ad aiutare i sordi profondi a capire meglio la prosodia e informazioni circa il senso della frase (per esempio distinguere una frase interrogativa da una affermativa) (Weisenberger et al., 1991). La maggior parte dei dispositivi a canale singolo riesce ad affrontare meglio il problema della scarsa sensibilità della pelle e dell'estrapolazione delle frequenze alte e veicolano per lo più le basse frequenze in modo costante con un'ampiezza che viene modulata seguendo la struttura del linguaggio (si vedano ad esempio Beguesse, 1976, Spens and Plant, 1983, Flanklin, 1984). Ovviamente, di contro, il problema di questi dispositivi è che vengono perse le informazioni spettrali dettagliate, mentre si assicura che il profilo intensità-tempo e le caratteristiche prosodiche del segnale vocale possano essere perepiti distintamente.

Broadstone et al (1988), valutando l'efficacia degli ausili a canale singolo e a due canali hanno concluso che l'aggiunta di un secondo canale non aumenta in modo significativo la capacità dei soggetti di identificare fonemi o vocali, eccetto nei compiti appositamente progettati per la rilevazione delle fricative.

Uno dei primi dispositivi a canale singolo, come accennato, è il tubo vibratorio di Gault (1924), che riusciva a veicolare principalmente energia a bassa frequenza (da 200 a 300 Hz), che riflette l'ampiezza generale dell'informazione linguistica. Con questo dispositivo i soggetti sono stati in grado di identificare correttamente ogni 1 di 120 frasi tratte dal discorso colloquiale dopo circa 100 ore di pratica. Le singole parole erano identificabili (da un insieme di 58) e si suppone che i soggetti potrebbero arrivare a discriminare l'uno dall'altro singoli suoni vocali. Similmente, il dispositivo di Beguesse (1976) veicolava frequenza di 250 Hz.

Grant (1980) ha utilizzato un dispositivo con 10 elettrodi in cui la frequenza fondamentale della voce, F0 (ad ampiezza costante), è stata 'suddivisa' spazialmente in diversi siti disposti lungo l'avambraccio degli utenti (per facilitare il rilevamento dei cambiamenti della frequenza fondamentale). L'autore riscontra un modesto guadagno di informazioni per due soggetti ed ipotizzò che quindi occorre veicolare la modulazione sia dell'ampiezza che della frequenza di Fo. Analizzando la velocità di linguaggio e il monitoraggio del discorso (De Filippo e Scott, 1978), l'autore riscontra che sia un udente reso artificialmente sordo, che un sordo profondo, dopo 20 ore hanno ottenuto un incoraggiante miglioramento e riuscivano a discriminare intonazione, ritmo, accenti della frase, difficili da discriminare senza il dispositivo.

Goldstein et al. (1983) riportano che un bambino sordo di 3 anni, che usava un apparecchio acustico, ma aveva solo un vocabolario a cinque parole della lingua dei segni, ha aumentato notevolmente la sua attenzione comunicativa quando utilizzava un dispositivo a canale singolo e ha poi acquisito un vocabolario di 400 parole in 7 mesi, grazie alla comprensione del ritmo del linguaggio.

Anche i 15 soggetti sordi nello studio di Spens e Plant (1983) dopo un periodo di pratica da 3 a 20 mesi (anche se 6 soggetti hanno ricevuto una formazione speciale per 2 ore settimanali per 12 settimane) hanno mostrato un significativo miglioramento nel test di monitoraggio del discorso (De Filippo e Scott, 1978), del riconoscimento dei suoni ambientali e un miglioramento anche nella lettura del labiale (n.b. Nel monitoraggio del discorso, gli autori presentano 5 con e 5 minuti senza aiuto dello sperimentatore). I soggetti hanno riportato più di 10 parole al minuto in 6 sessioni di allenamento e la percentuale di parole corrette è migliorata dal 35% al 60%.

In lavori successivi, Plant et al. (1984a) riportano significativi miglioramenti nella lettura delle labbra a livello fonemico, della comprensione di parole e frasi e nei punteggi di monitoraggio del discorso, sempre quando le informazioni della frequenza fondamentale F0, venivano presentate tattilmente (dispositivo vibratorio). Nel 1986, Plant ha testato il livello di capacità di discriminazione di consonanti di quattro soggetti sordi, solo con lettura labiale e lettura labiale e dispositivo a canale singolo. L'autore notò un miglioramento sia della percezione delle consonanti sia nel modo di articolazione, anche in discorsi. Gli autori interpretano questo risultato come evidenza che fornendo questo tipo di informazione dovrebbe aumentare notevolmente la quantità di informazioni segmentali disponibili. Allo stesso modo, il cambiamento fondamentale della frequenza serve anche a segnalare caratteristiche prosodiche come gli accenti e l'enfasi.

Risultati meno impressionanti sui dispositivi a canale singolo e due canali sono stati riportati da Weisenberger e Russell (due canali, REF) che hanno riportato un modesto riconoscimento di vocali e scarso riconoscimento delle vocali in set chiusi, Miyamoto et al. (canale singolo, REF) hanno riportato scarsi risultati nel monitoraggio del discorso (5 parole al minuto) e Bernstein et al. (REF) hanno riportato uno scarso miglioramento per l'intonazione e accenti (solo del 3% rispetto ad una condizione di sola lettura delle labbra).

I dispositivi multicanale presentano il vantaggio di veicolare informazione spettrale. Di contro, il loro utilizzo è più complesso, sia perché sono più difficili da codificare, sia per la loro scarsa portabilità (almeno in passato). Un argomento che è stato portato a loro favore, è che siano più simili e che simulino, gli impianti cocleari, in termini di quantità di informazione veicolata.

Risultati a favore dei dispositivi multicanale sono stati riportati da Engelmann e Rosov nel 1975 in uno studio di apprendimento di parole e frasi con soggetti adulti udenti e bambini sordi che hanno usato un dispositivo vibrotattile a 25 canali. Utilizzando solo l'informazione tattile, venivano presentate inizialmente solo 5 parole e man mano sono state aggiunte altre parole. Il miglior soggetto (8 anni) dopo 34 settimane di allenamento ha imparato 60 parole (di un pool di 75 parole) con l'80% di accuratezza per poi accelerare i ritmi di apprendimento e arrivare, in altre 13 settimane di allenamento, a 122 parole (di un pool di 135 parole) con il 90% di accuratezza. Un ragazzo sordo, in una condizione in cui veniva presentata la stimolazione tattile e apparecchio acustico, ha imparato in 8 settimane, 35 parole (94% di accuratezza). Se si considera che lo stesso ragazzo raggiungeva solo con l'apparecchio acustico il 65% di accuratezza, e solo con l'informazione tattile il 40%, gli autori pensano che integrava bene le due informazioni (dopo 21 settimane ha imparato 60 parole con il 90% di accuratezza, e dopo 26 settimane 100 parole con il 78% di accuratezza usando entrambe le modalità). Come accennato nel paragrafo del training, gli autori hanno concluso che sono necessarie centinaia di ripetizioni corrette per imparare le semplici discriminazioni tattili, ma che, con buone condizioni di apprendimento, il tasso di apprendimento accelera una volta acquisita una serie iniziale di 30-40 parole.

Nello studio di Brooks e Frost nel 1983 con soggetti udenti, utilizzando un vocoder a 16 canali (sessioni giornaliere di circa mezz'ora per 5 giorni alla settimana), un soggetto, dopo 55 ore di allenamento ha imparato 150 parole e in ulteriori 25.5 ore di allenamento è arrivato ad un totale di 250 parole all'80% di accuratezza. Nel test di monitoraggio del significato di un discorso discorso collegato, il soggetto ha raggiunto un tasso di 51 parole corrette al minuto quando utilizzava il dispositivo multicanale.

Friel-Patti e Roeser nel 1983 eseguono uno studio con quattro bambini sordi profondi che hanno utilizzato un dispositivo a tre canali in classe e durante la terapia (una volta ogni tre settimane per mezz'ora). Gli autori hanno confrontato un periodo di sei mesi in cui i bambini hanno indossato il dispositivo e apparecchio acustico per 10-11 ore a settimana per 16 settimane, con un periodo di 6 mesi in cui non hanno utilizzato il dispositivo.
Analizzando durata, contenuto e qualità della produzione della comunicazione (lingua dei segni e parlata), gli autori hanno riportato nei mesi in cui i bambini hanno usato il dispositivo e l'apparecchio, un miglioramento della durata media della produzione comunicativa da una parola in dieci minuti, a una parola e mezza in quattro minuti, mentre quando non hanno usato il dispositivo tattile la durata della comunicazione è diminuita nel semestre da una media di circa 3 minuti a circa 2 minuti, di comunicazione efficace (in 10 minuti di discorso). Concludono quindi che l'utilizzo del dispositivo è stato importante non solo per la durata, ma anche per la qualità e capacità di vocalizzazione dei bambini (facilitata nella sua elicitazione).
Nello studio di Lynch et al. (1988), due adulti sordi profondi congeniti hanno imparato ad usare un dispositivo elettrocutaneo a 16 canali (Tacticon 1600). I soggetti hanno imparato a identificare 50 parole durante 47 ore (soggetto uno) e 41 ore (soggetto due) di addestramento, con un tasso di successo del 41,6% in tutte le sessioni.
Risultati incoraggianti a favore dei dispositivi multicanale sono quelli riportati dal gruppo di Clark che ha utilizzato un dispositivo multicanale elettrotattile (8 canali) chiamato Tickle Talker (Blamey e Clark, 1987), utilizzato principalmente come dispositivo supplementare, oltre alle informazioni uditive residue e alle abilità di lettura delle labbra. Nell'articolo di Galvin et al. (1993) viene riportata una sintesi utile per capire come è fatto e la tipologia di informazione veicolate (si veda fig.1 in appendice). Da un punto di vista psicologico, questa sintesi è utile per capire cosa effettivamente veicolano, in termini di percezione linguistica percepita, le varie informazioni presente nello spettro linguistico (frequenza, ampiezza, ecc).

L'estrazione del segnale del parlato, può quindi utilizzare due diversi sistemi o 'strategie', come definite da Loizou (1998): l'utilizzo delle onde ('the waveform strategies'), che presenza alcuni tipo di onde sonore (in forma analogica o pulsatile) trasferite tramite il filtraggio del segnale in diverse bande; o l'estrazione di caratteristiche ('the feature-extraction strategies') che prova a trasferire caratteristiche spettrali, come le formanti, tramite algoritmi di estrazione di caratteristiche (si veda fig.2 in appendice).

Alcuni studi hanno confrontato l'utilizzo di dispositivi a canale singolo vs. multiplo, ma come accennato in precedenza, vi sono risultati contrastanti in merito alla loro efficacia, o meglio, legati al tipo di compito utilizzato negli studi e al tipo di utenza.

Pickett e McFarland (1985) hanno ritrovato per frasi semplici non vi è differenza tra dispositivi a canale singolo e multicanale e che c'è una grande variazione tra i soggetti (valutano soggetti con impianto cocleare).

Weisenberg et al. (1991) hanno confrontato le prestazioni ottenute utilizzando due dispositivi a singolo canale e due dispositivi multicanale. Gli autori hanno riscontrato performance simili per diversi compiti di analisi linguistica, ma per compiti di identificazione di informazioni fonetiche fini, le performance erano migliori per i dispositivi multicanale. Nel loro studio, quindi, non hanno confermato la superiorità dei dispositivi a canale singolo per il ritmo sillaba e gli accenti. Anche nel confronto del monitoraggio del discorso, i miglioramenti nel compito sono stati da 10 a 15 parole per minuto per il dispositivo a due canali e da 40 a 50 parole al minuto per il dispositivo a 16 canali, rispetto alla sola lettura delle labbra.

Di contro, nel confronto tra dispositivi vibrotattili a canale singolo vs. multicanale (24 canali), Carney e Beachler (REF) hanno riscontrato che nei compiti che comportano il riconoscimento di funzioni suprasegmentali, come per il numero di sillabe, sillabismi e intonazione, il dispositivo a canale singolo è risultato essere significativamente migliore rispetto al multicanale.

Nel confronto tra dispositivi a due canali e 16 canali, Lynch et al. (REF) hanno trovato risultati uguali o superiori per il dispositivo a due canali per il monitoraggio del discorso, ragion per cui gli autori concludono che multicanale non fornisce più informazioni di quelle disponibili con il dispositivo a due canali, almeno per questo tipo di compito.

Riguardo all'area in cui è stata applicata la stimolazione, sono state stimolate sia la schiena, che le dita, lingua (Hueber et al, 2010), avambraccio (Grant, 1980), sul petto (Goldstein et al., 1983), sulla pancia (cintura di Friel-Patti e Roeser, 1983). In questa sede si farà riferimento principalmente alle stimolazioni delle dita della mano, dal momento che i risultati più convincenti sono risultati proprio da questo tipo di stimolazione (e.g. Clark et al, REF). In aggiunta, si noti che la stimolazione del dorso ha mostrato risultati molto soddisfacenti nella sostituzione sensoriale visuo-tattile (REF). Inoltre, sebbene la stimolazione della lingua sia stata utilizzata, non sembra essere né piacevole, né utilizzabile all'atto pratico nella quotidianità.

Riguardo alla tipologia di stimolazione, come accennato, sono state usate stimolazioni di tipo vibratorio ed elettrico, ma essi presentano il maggiore ostacolo della perdita di sensibilità della pelle dopo qualche ora dal loro utilizzo, oltre al fatto che il tipo di stimolazione non risultava piacevole (soprattutto per i dispositivi elettrotattili, Kirman 2006).

Dispositivi tattili per la comprensione del linguaggio:il superamento dei limiti.

Come già detto, i dispositivi tattili usati fino ad oggi per veicolare informazioni linguistiche avevano alcuni limiti, tra cui: la loro scarsa portabilità (erano ingombranti), scarsa piacevolezza (soprattutto quelli elettrotattili), difficoltà nella decodifica delle informazioni veicolate (dovuta anche alla desensibilizzazione della pelle dopo un breve uso), grande sforzo e lunghi periodi di addestramento.

Per risolvere il primo limite, la scarsa portabilità, la protesi *FeelSpeech* è stata resa portabile. Tutte le versioni che sono state costruite richiedono l'utilizzo di un computer; mentre il dispositivo presenta: un 'decodificatore' (una 'scatola di splastica dalle dimensioni di 12 cm x 17 cm), cavi (da 1 a 4 cavi) ed una parte terminale che viene tenuta come 'joystick' , della grandezza di un 10 cm x 1 cm (si veda fig.).

Riguardo al secondo limite, la scarsa piacevolezza, e il terzo, la difficoltà di decodifica dell'informazione tattile legata alla desensibilizzazione della pelle, la protesi *FeelSpeech* trasmette le informazioni del parlato in modo completamente differente rispetto alla vibrazione o alla stimolazione elettrica. La stimolazione tattile, presenta un movimento di 'bottoni' di plastica, che abbiamo chiamato sistema di *microswiping*. Rispetto alle stimolazioni elettriche o vibratorie, il *microswiping* è piacevole. Inoltre, potenzialmente è in grado di fornire molte più caratteristiche del parlato (es. fonemi) visto che i pattern di movimento sono rilevabili molto facilmente.

Riguardo al terzo limite, ovvero il grande sforzo e i lunghi periodi di addestramento richiesti, si è cercato di svolgere contributi sperimentali che potessero confermare la sua facilità ed efficacia di utilizzo.

Sono state utilizzate diverse versioni del dispositivo (si veda parte sperimentale per i dettagli dei dispositivi utilizzati per ogni esperimento).

Una versione semplice del dispositivo, prevede una matrice 4x4 di bottoni, tenuta dai partecipanti come un joystick con una mano, e ricevono la stimolazione sul pollice. Questa versione è stata utilizzata nell'esperimento dell'audiolibro e presenta ai soggetti l'ampiezza generale del linguaggio. L'andamento dell'ampiezza viene percepito in maniera propozionale all'area stimolata (l'ampiezza maggiore comporta tutti i bottoni alzati, ovvero l'intera superficie del pollice viene stimolata).

Una versione più complessa del dispositivo, utilizzando sempre una matrice 4x4 di bottoni, e con stimolazione sul pollice, veicola un gran numero di caratteristiche (fonemi). Il *microswiping*, presenta dei movimenti in 12 direzioni: sinistra-in su, in basso a destra, in alto a destra, in basso a destra -left, up-right, down-left, down-right e quattro direzioni diagonali. La durata totale di ciascuna di queste stimolazioni in movimento è nell'intervallo 50-200 ms, a

seconda del particolare modello. Ne è stato fatto utilizzo negli esperimenti di discriminazione BA-VA, nel pilota McGurk audio-tattile e nel pilota con linguaggio tattile e pilota con l'audiolibro.

E' stata costruita una versione con un approccio simile ai vocoders. In questo caso sono utilizzate 5 matrici (ciascuna per un diverso dito) con matrice 2x4 e fornisce una stimolazione tattile proporzionale alla quantità di energia in ciascuna banda (si veda descrizione nella parte sperimentale). Abbiamo anche usato questo sistema per l'esperimento della finestra temporale e nel pilota solo tattile e tattile con lettura delle labbra.

Dispositivi tattili per la comprensione del linguaggio: focus sul training.

Engelmann e Rosov (1974) suppongono che se i dispositivi tattili fossero stati adeguati (ovvero riuscissero a fornire le informazioni necessarie per un'adeguata percezione del linguaggio), gli utenti avrebbero dovuto imparare ad usarli rapidamente, se non immediatamente. Inoltre, affermano che per massimizzare i loro potenziali benefici, l'utilizzo degli strumenti tattili deve essere accompagnato da un programma formativo efficace e adattivo.

Tra i fattori importanti per lo sviluppo di un programma di formazione per l'uso con dispositivi tattili Galvin et al. (1993) citano: 1) il tipo di utenza, 2) la motivazione e l'uso del dispositivo, 3) il tipo di informazioni presentate dal dispositivo tattile, 3) quantità del training, 4) il tipo di risposte e di test impiegate nel training, 5) formati di risposta utilizzati nella formazione e procedure di valutazione. Di seguito verranno analizzati i punti citati.

Utenza. Il training dovrebbe essere flessibile per garantire la sua applicabilità ad un'ampia popolazione. Esistono diverse importanti caratteristiche dell'utente da tenere in conto: l'età, le capacità uditive residue, la conoscenza formale della lingua, l'abilità di comunicazione in generale, la capacità di integrare le informazioni vocali provenienti dalle diverse modalità sensoriali e le aspettative sull'efficacia del dispositivo .

Il progetto *FeelSpeech* è rivolto ai pazienti con problemi di udito che hanno avuto precedenti esperienze di udito (ad esempio problemi di udito comune delle persone anziane).

Motivazione e utilizzo del dispositivo. La motivazione per l'utilizzo del dispositivo è strettamente legata all'ambiente intorno al soggetto ed è fondamentale per migliorare, in termini di tempo e prestazioni, l'uso del dispositivo. Questo è importante soprattutto per i bambini, e di conseguenza, per gli adulti (genitori e insegnanti) che possono giocare un ruolo fondamentale nel motivarli ad utilizzare il dispositivo nei diversi contesti e in diverse situazioni della vita quotidiana(familiare e scolastico). D'altra parte, ciò implica indirettamente che il training deve avere come requisito la sua semplicità, ovvere deve poter essere facilmente

somministrabile da chi non ha alcuna formazione o struttura specializzata. In tal senso, ci si riallaccia al primo punto sulla facilità e flessibilità di utilizzo per una vasta gamma di utenti. L'addestramento in diversi contesti dovrebbe facilitare la generalizzazione dell'utilizzo del dispositivo ad ulteriori contesti, non addestrati originariamente.

Tipo di informazioni presentate dal dispositivo tattile. Il training va programmato in base al tipo di informazioni che vogliamo trasmettere, nonché le informazioni del discorso veicolabili tramite un dispositivo nello specifico e la salienza di tali informazioni per la popolazione di utenti scelta. Come abbiamo detto (vedere paragrafi precedenti), i dispositivi differiscono sia dalle informazioni di discorso forniti e dal modo in cui vengono presentate all'utente, ciascuno con la propria efficacia in base a determinate caratteristiche della lingua (ad esempio, come si è visto, i dispositivi a uno o due canali sono migliori nel fornire le frequenze alte, utili a discriminare le informazioni prosodiche e ritmiche del parlato).

Quantità e durata del training. In letteratura gli studi presentano diversi tempi di addestramento, differenti soprattutto in base agli obiettivi degli studi (per esempio a seconda se si tratti di identificare e discriminare correttamente fonemi, parole e comprensione di conversazioni). Vengono di seguito riportati alcuni esempi.

Oiler et al. (1980) hanno addestrato adolescenti sordi per discriminare alcune coppie di parole difficili da leggere per circa un'ora di pratica usando un dispositivo a 24 canali. Gli autori hanno verificato che alcuni suoni del discorso precedentemente indistinguibili dalle persone sorde sono stati immediatamente riconosciuti; mentre altri suoni del discorso hanno bisogno di tempi maggiori di training. Questo lavoro suggerisce che dopo una pratica minima, alcune coppie di parole possono essere discriminate con successo, ma per comprendere correttamente il discorso è necessaria una pratica lunga ed esperienza .

Alcuni studi concordano circa il fatto che una pratica individuale con esperti potrebbe variare da 20-50 a 70-80 ore, diluite nell'arco di un anno o più. Per esempio, nel già citato studio di Engelmann e Rosov (1974) il periodo di training dei con soggetti normoudenti è durato quasi un anno (11 mesi, in dettaglio i soggetti hanno svolto 80, 70, 50, 20 ore). Le sessioni di allenamento hanno durato da 20 a 60 minuti (di solito 30 minuti) e la loro formazione con soggetti sordi ha durato un tempo supplementare (un ulteriore periodo di formazione, inizialmente di un'ora al giorno per sei giorni alla settimana e successivamente per cinque giorni alla settimana). In uno studio successivo (1975) hanno somministrato ai loro soggetti (sia sordi che udenti resi sordi artificialmente) un training da 20 a 22.5 ore di formazione con un dispositivo con stimolazione vibrotattile e hanno verificato che i loro soggetti sono riusciti

a discriminare da 15 a 152 parole (liste di parole simili, diverse solo per un fonema da una, due, tre e quattro sillabe).

Un altro esempio di studi con lungo periodo di training è lo studio di acquisizione vocabolario svolto da Brooks e coll. (1985) . Il soggetto con prestazione di apprendimento migliore tra due udenti resi artificialmente sordi (nello studio di Scilley et al. REF ha imparato a riconoscere 150 parole in 55 ore di pratica, mentre l'altro 70 parole in 40.5 ore) è arrivato a identificare complessivamente un vocabolario a 250 parole dopo 80.5 ore di allenamento. Il risultato più importante dello studio di Brooks et al. (1985) riguarda il fatto che hanno dimostrato che le conoscenze tattili acquisite nella formazione di un vocabolario possono essere generalizzate all'identificazione di nuove parole e che quindi, le prime fasi di apprendimento sono più lente mentre in periodi successivi l'apprendimento di nuove parole richiede minor tempo. Probabilmente, ciò è legato al fatto che inizialmente i soggetti devono familiarizzare con il dispositivo, ma che una volta capita la 'legge' di decodifica, l'apprendimento risulta più facile. Blarney et al. (1988) ipotizzano che le 70 ore di pratica fornita a soggetti normalmente udenti nel loro studio non era sufficiente per massimizzare i vantaggi del dispositivo in riconoscimento di parole e quindi di fatto il vantaggio dell'utilizzo dei dispositivi tattili non può essere stimato.

Cowan et al. (1991) riportano un miglioramento della percezione del discorso in due soggetti con deficit uditivi durante un training di 28 e 43 sessioni settimanali.

In generale, Weisenberger (1987) riporta un fattore molto importante riguarda la frequenza, ovvero che l'addestramento quotidiano comporta livelli di prestazione più elevati. Diversi studi (specialmente studi sui bambini) riportano la combinazione di una formazione individuale (ogni sessione di allenamento per meno di 40 minuti) più la formazione quotidiana a casa. Alcantara et al. (1990) riportano che se non è possibile una sessione giornaliera (ad esempio, a scuola, per i bambini), la formazione dovrebbe essere limitata al raggiungimento di obiettivi quali le abilità percettive di base e di integrazione, che l'utente può mantenere anche dopo la pratica e generalizzare in contesti non addestrati.

Anche Galvin et al. (1993) suggeriscono che una pratica ideale dovrebbe includere una sessione quotidiana di pratica più la formazione individuale, almeno due volte alla settimana, con uno specialista. Gli autori suggeriscono di guardare alla quantità di tempo per la pratica, che dovrebbe essere considerata in termini di costi.

Tipologia di training La letteratura suggerisce due tipi principali di attività di training, che differiscono per il tipo di obiettivo, ovvero il tipo di analisi e il livello di comprensione del discorso: l'analitico e il sintetico. Le attività di training analitico si concentrano sulla

discriminazione delle funzioni di discorso (cioè la discriminazione fonetica), mentre i compiti di training sintetico si riferiscono al all'accuratezza e comprensione di conversazioni quotidiane. Come accennato, una buona formazione dovrebbe consentire ai partecipanti di generalizzare le abilità linguistiche da formate a parole non addestrate. Il training di tipo sintetico, focalizzato sulla comprensione generale del discorso, dovrebbe poter consentire all'utente, durante una conversazione, di sottolineare l'importanza del contesto nella percezione delle frasi. L'obiettivo di introdurre le parole in un contesto 'sintetico' risulta essere importante, dal momento che in letteratura viene riportato che riferisce che i soggetti sordi o con deficit uditivi, presentano dei deficit in generale, nella comprensione del nel senso "sintattico" (Engelmann e Rosov, 1974).

Cowan et al. (1988) hanno addestrato quattro soggetti con deficit uditivi e sette soggetti normo udenti (resi artificialmente sordi). I soggetti udenti hanno avuto 35 ore di training analitico e 35 in analitico più sintetico, mentre i soggetti con deficit uditivi hanno svolto un training combinato. Per il training sintetico hanno usato la tecnica suggerita da De Filippo e Scott (1978) in cui gli utenti devono ripetere ciò che viene detto dallo sperimentatore, e per il training analitico e sintetico hanno somministrato prove di identificazione progressiva di vocali, le consonanti, le parole e le frasi sono state usate, in combinazione con la tecnica di De Filippo e Scott (1978). Il training sintetico è stato somministrato nelle condizioni di sola lettura delle labbra e lettura delle labbra più stimolazione tattile. Per il training analitico è stata prevista anche una condizione di stimolazione solo tattile. Gli autori hanno riportato nessuna differenza in base al tipo di training, visto che per entrambe venivano riconosciute 55 parole per minuto nella condizione tattile più lettura delle labbra (la sola lettura delle labbra comporta il riconoscimento di 36 parole al minuto) per il gruppo dei normoudenti (dopo 70 ore di training). Il gruppo dei soggetti con deficit uditivi in entrambi i casi venivano riconosciute 37 parole per minuto nella condizione tattile più lettura delle labbra (la sola lettura delle labbra comporta il riconoscimento di 24 parole al minuto (dopo le 35 ore di allenamento). Gli autori riportano che indipendentemente dal tipo di training, tutti i soggetti hanno mostrato un rilevante miglioramento nella comprensione di parole al minuto.

Dal momento che entrambi i tipi di training sono fondamentali, Galvin et al. (1993) suggeriscono che un programma di formazione dovrebbe includere compiti appartenenti ad entrambe queste due tipologie (ricezione di fonemi e comprensione del discorso). Un modo per utilizzare entrambe le tipologie di training, presente in letteratura, è quello di presentare compiti di entrambe le tipologie, in ordine gerarchico, generalmente partendo dal tipo analitico a quello sintetico. In altre parole, si passa dalla discriminazione di caratteristiche alla conversazione,

nonché dai compiti più facili a quelli più difficili. Di solito, quindi, una sessione di training inizia con compiti di discriminazione di caratteristiche linguistiche (ad esempio discriminazione fonetica di vocali o consonanti), seguita da parole isolate, e dopo vengono presentate frasi intere e piccoli discorsi (si veda Galvin e altri, 1993, Alcantara et al, 1990). I compiti di discriminazione fonetica di solito sono costituiti da elenchi di parole semplici, mentre il discorso viene formato dall'unione di frasi che presentano le medesime parole del training, o parole nuove. Una procedura già citata che viene utilizzata per il livello conversazionale è quella del '*tracking procedure*', letteralmente 'procedura di monitoraggio del linguaggio' di De Filippo e Scott, (1978) in cui l'utente deve ripetere le frasi del parlante (si veda la sessione sui compiti per i dettagli). Engelmann e Rosov (1974) hanno utilizzato anche compiti che includevano la prosodia: inflessione, ritmo, azione implicita del linguaggio. Gli autori gestivano la quantità di tempo per ogni passo di questa procedura gerarchica in base alle capacità discorsive dei soggetti; e la proporzione del tempo trascorso per ciascuna frase e le attività di conversazione aumentavano in base alle competenze e l'esperienza acquisita dai soggetti. Si noti infatti, che mentre nella maggior parte degli studi questa gerarchia dei compiti dei due tipi di training era 'sequenziale', a volte diluita nell'arco di diversi mesi (per esempio i soggetti fanno solo il compito di parola isolata finché non raggiungono il 70% della precisione e quindi possono iniziare con i compiti della frase, si veda REF), Engelman e Rosov (1974) hanno distribuito i compiti diversamento nel tempo, presentando tutti i compiti (dal più semplice al più complesso) in ogni sessione e dividevano l'ora in modo proporzionale alle abilità linguistiche acquisite dai soggetti nel momento della sessione. In altre parole, mentre in una fase iniziale dedicavano più tempo ai compiti di discriminazione delle caratteristiche e meno tempo ai compiti di conversazione, alla fine del periodo di training, quando i soggetti erano più esperti, hanno dedicato più tempo ai compiti più complessi conversazionali e meno tempo per le attività di base.

Per quanto riguarda le 'condizioni' utilizzate, o il modo in cui vengono presentati gli stimoli, in genere prese isolatamente o combinate, si possono avere le condizioni tattile e tattile più la lettura delle labbra e negli studi con soggetti con parziale deficit uditivo, può essere presente anche la modalità uditiva (con le capacità residue). Alcuni studi riportano tutti questi tre tipi di formazione (ad es. Cowan et al., 1988, Cowan et al. 1990, Galvin et al 1993, Sarant et al., 1996, Alcantara et al., 1993) e altri solo la formazione con la modalità tattile (Lynch et al., 1988; Engelmann e Rosov, 1974; Weisenberger e Broadstone, 1988).

Il training che prevede una condizione "solo tattile" implica che i soggetti non guardano il parlante (sperimentatore), quindi ricevono solo le informazioni del discorso o parole isolate,

solo attraverso l'aiuto del dispositivo tattile. Ovviamente, i partecipanti udenti vengono resi artificialmente sordi, in genere indossando cuffie attraverso le quali viene trasmesso il rumore bianco. Uno dei motivi per privilegiare questo tipo di training solo tattile può essere motivazionale, ovvero far rendere consapevoli gli utenti che la percezione tattile non è difficile (Galvin et al., 1993). Inoltre, la condizione "solo tattile" è un buon approccio per capire come funziona il dispositivo e prenderne buona familiarità. In aggiunta, la maggior parte degli studi ha ipotizzato che la comprensione del linguaggio sia più facile se i partecipanti utilizzano più di una modalità sensoriale, quindi se i partecipanti possono eseguire l'attività solo con il dispositivo tattile, lo faranno con migliori prestazioni con l'aggiunta delle altre modalità sensoriali. Infine, training tattili che inizialmente vengono applicati a parole prese singolarmente, hanno mostrato prestazioni superiori in studi di monitoraggio di discorsi (Weisenberger, 1987). Se si utilizza una procedura gerarchica o negli studi che utilizzano più di una condizione, la condizione ‘solo tattile’ coincide generalmente con compiti semplici e altamente strutturati, come il rilevamento vocale, l'ampiezza e il numero di sillabe e la discriminazione dello stress (Blarney et al., 1988; Cowan et al, 1989b).

La condizione ‘tattile e lettura delle labbra’, ovvero, l’utilizzo della lettura labiale in simultanea alla stimolazione tattile nel training, è presente in diversi studi. Gli studi con adulti udenti e con deficit uditivi hanno evidenziato che una combinazione sistematica di training di lettura labiale e tattile migliora la capacità dei soggetti di utilizzare l'input tattile (Cowan et al 1988, Galvin et al 1993, Cowan et al., 1989a).

La condizione ‘tattile e uditiva’ consiste nella compresenza di utilizzo dell’apparecchio acustico e dispositivo tattile. Negli studi che utilizzano una procedura gerarchica e una formazione in più condizione, la condizione ‘tattile e uditiva" può essere applicata sia alle abilità di percezione linguistica (Alcantara et al., 1990b) sia alla discriminazione delle parole (Cowan et al., 1990).

Nella condizione ‘tattile, uditiva e lettura delle labbra’ i partecipanti possono guardare il parlante, indossare l'apparecchio acustico e usare il dispositivo tattile. Di solito questa condizione viene utilizzata per la discriminazione del livello della frase o il discorso conversazionale, in particolare per i partecipanti sordi e con deficit uditivi.

Tipologia di test per il training.

Dal momento che l’obiettivo del training è quello di consentire una familiarizzazione con il dispositivo tattile, che comporti un’agevolazione nella comprensione del discorso, i test utilizzati devono riflettere (ed essere quindi simili) ai contesti di una conversazione normale. In letteratura vengono riportati due tipi di formati di risposta: chiusi e aperti. Il formato di

risposta chiuso si riferisce a una scelta forzata tra diverse opzioni. Questo formato di risposta è simile a quando si ascolta una parola, ma si ha dubbi su diverse opzioni. I compiti che richiedono questo tipo di formato di risposta presentano minimo 2 e massimo 5 opzioni di risposta, che sono simili al target da individuare (in generale parole o frasi). Per i bambini, le le opzioni di risposta possono essere delle immagini (si veda Cowan et al., 1990). Il formato di risposta aperto non fornisce alcuna opzione al partecipante. Questo formato di risposta richiede abilità linguistiche più sofisticate ed è utile per sviluppare abilità nella comprensione del discorso.

I soggetti addestrati con compiti chiusi hanno mostrato un miglioramento generale in tutti i compiti a prove chiuse, mentre chi viene allenato nelle prove aperte sviluppa abilità specifiche nei compiti di tipo aperto (Alcantara et al., 1990b; Bode & Oyer, 1970).

Inoltre, gli studi con gli adulti che utilizzano Tickle Talker (Alcantara et al., 1990a e Cowan et al., 1989a) hanno mostrato che un training con compiti con risposte chiuse può essere fondamentale per familiarizzare inizialmente con un dispositivo tattile, e che l'utilizzo dei compiti aperti erano necessari per incoraggiare e facilitare la percezione e l'integrazione delle parole nel discorso. Si può quindi immaginare, che per sviluppare a pieno entrambe le due capacità (di discriminazione di dettagli e di comprensione del discorso) sia meglio utilizzare entrambe le tipologie di prove. Anche Bode e Oyer (1970) hanno sostenuto l'inclusione di entrambi i formati di risposta, dal momento che percezione quotidiana del parlato comprende sia decisioni chiuse che aperte.

Durante la formazione iniziale, potrebbe essere utile fornire un feedback ai soggetti. Ad esempio, Lynch et al. (1988) presentavano gli stimoli solo tattilmente e il soggetto doveva ripetere che parola era stata detta (sperimentatore e soggetto erano seduti schiena contro schiena). Gli autori presentavano uno stimolo (X) fino a quando il soggetto rispondeva e se la risposta del soggetto (Y) er corretta, lo sperimentatore dava un feedback positivo "sì" e presentava lo stimolo successivo; mentre, se la risposta del soggetto era errata, lo sperimentatore dava il feedback negativo e ripeteva lo stimolo precedente rinforzando l'associazione tattile corrispondente allo stimolo target (esempio "no Y, X ").

Procedure di valutazione. Le procedure di valutazione sono importanti per valutare l'efficacia del training e del dispositivo tattile. Le procedure di valutazione possono essere eseguite in prima, durante e dopo il periodo di training. Le procedure di valutazione prima del training hanno come obiettivo quello di verificare le abilità discorsive e il livello di comprensione dei soggetti. Le procedure di valutazione durante la formazione mirano a monitorare i progressi

dei soggetti e di verificare l'efficacia del dispositivo tattile. Le procedure di valutazione dopo il periodo di training mirano a verificare gli effetti a lungo termine, inclusa il mantenimento del tipo di stimolazione nel tempo (per esempio, quanto può essere utilizzato il dispositivo prima di una desensibilizzazione), l'estensione dell'utilizzo del dispositivo in base alla sua piacevolezza e praticità di utilizzo in vari contesti e la possibilità di generalizzare la formazione alla presentazione di nuove parole (capacità di generalizzazione).

Si noti che uno dei limiti delle procedure di valutazione del training per l'utilizzo dei dispositivi tattili è che la prestazione misurata prima del training non necessariamente riflette i vantaggi della percezione del linguaggio ottenuti dopo il training, ma la familiarizzazione con i test utilizzati (Engelmann & Rosov, 1975; Weisenberger, 1987; Weisenberger & Miller, 1987).

Test e batterie di test. Di seguito vengono riportati alcuni dei test più influenti nella valutazione dei dispositivi tattili, utilizzati sia per il training che in fase di test finale di valutazione dell'efficacia dei dispositivi. Come citato, i test mirano a valutare le capacità di discriminazione di caratteristiche (funzioni del discorso) e comprensione di conversazioni. Alcuni test sono inizialmente stati usati per valutare l'efficacia di tecniche riabilitative, non necessariamente legate all'utilizzo di dispositivi tattili per la comprensione linguistica.

In tabella vengono riportati solo i test utilizzati per la popolazione adulta (per alcuni di essi c'è anche una versione per bambini).

Nome del test/batteria	**Descrizione**
The Speech, Spatial, and Qualities of Hearing Scale (SSQ; Gatehouse and Noble, 2004) SSQ-B version (Jensen et al, 2009)	Valuta le difficoltà uditive in circostanze di ascolto dinamiche e complesse (per esempio con rumore o più parlanti). La versione SSQ-B è stata creata appositamente per valutare i benefici degli apparecchi acustici.
Speech tracking (DeFilippo and Scott, 1978) (CDT; De Filippo & Scott,1978)	In questa procedura di monitoraggio della comprensione del discorso, lo sperimentatore legge ad alta voce una porzione di testo (in genere una storia) a una normale velocità di conversazione e un ascoltatore cerca di ripetere la frase. Se il destinatario non ripete la frase correttamente, lo sperimentatore ripete la parola corretta o cerca di aiutarlo utilizzando strategie come ripetere una parte o tutta la frase, rivede la frase (semplificandola) o o fornisce degli indizi sul contesto leggendo la frase o il periodo successivo. La misura del successo del monitoraggio è il grado di avanzamento attraverso la storia, misurato in numero di parole

	per minuto.
Central Institute for the Deaf (CID) everyday sentences (Tonnison, 1974)	Il test è composto da frasi di una potenziale conversazione quotidiana per parametri quali il vocabolario, la lunghezza della frase, la struttura sintattica e la ridondanza. Valuta le capacità comunicative ricettive.
Consonant Nucleus Consonant (CNC) words (Peterson and Lehiste, 1959, 1962) The Maryland CNC Test (GD Causey, 1984)	10 liste di 50 parole monosillabiche
NU-6 word lists (Tillman and Carhart, 1966)	50 parole monosilabiche
Minimum Speech Test Battery (MSTB) (2011)	Test con frasi registrate in contesti rumorosi. Viene usato principalmente per utenti di impianti cocleari.
Minimal Auditory Capabilities (MAC) battery (Owens et al., 1985)	Batteria completa per la comprensione del linguaggio a diversi livelli (ad esempio subtest per l'identificazione di vocali e consonanti)
Four Alternative Auditory Feature test (AFAAF) (Foster and Haggard, 1987)	Identificazione di parole presentate in una frase (set chiuso).
The Abbreviated Profile of Hearing Aid Benefit (APHAB) (Cox RM, Alexander GC, 1995)	Misura problemi dell'udito in contesti quotidiani
The Device-Oriented Subjective Outcome (DOSO) Scale (Cox et al. , in press)	Chiede all'ascoltatore di specificare come sente in diverse situazioni
Speech in Noise (SIN™) Test (Etymotic Research, 1993).	Presenta parole la cui presenza è altamente o poco probabile in una frase
Hearing in Noise Test (HINT) (Nilsson, Soli, & Sullivan, 1994)	Riconoscimento di frasi (set apert) in conversazioni con contesti di quiete o con rumore di fondo

Conclusioni sul training. L'analisi della letteratura conferma l'importanza di un training efficace e adattivo per massimizzare i benefici disponibili dei dispositivi tattili nella comprensione del discorso, sia per popolazione di sordi che per persone con deficit uditivi. Sono stati riportati i fattori più importanti da considerare nella progettazione di un programma di training: caratteristiche degli utenti, motivazione e utilizzo del dispositivo, informazioni

presentate dal dispositivo tattile, quantità di formazione, tipo di attività di formazione, tipi di formati di risposta, procedure di valutazione, test e batterie.
Dalla revisione della letteratura è emerso che un buon programma di training dovrebbe essere svolto una o due volte alla settimana con un esperto e usato quotidianamente e in contesti diversi, deve includere attività di addestramento analitico e sintetico, formati di risposta chiusi e aperti, e avere una sequenza gerarchica dai compiti di livello di discriminazione di caratteristiche al livello di conversazione.
Secondo la teoria della sostituzione sensoriale, una buon training dovrebbe riuscire a dare la sensazione di sentirsi meglio sostituendo le informazioni uditive mancanti con le informazioni tattili.

1.4 Costruzione del percetto: dalla percezione alla rappresentazione della realtà.

Come visto, la percezione, più volte citata, è l'insieme dei processi per mezzo dei quali riconosciamo, organizziamo e diamo un senso alle sensazioni relative alle molteplicità degli stimoli ambientali. Nel presente lavoro, si è revisionata la letteratura in merito ai processi di percezione e, in particolare, del risultato percettivo di fenomeni quali l'integrazione di più stimoli, all'interno delle singole modalità, uditiva e tattile, e audio-tattile. Si è visto che il

risultato del processo di integrazione, e l'integrazione stessa, implica una costruzione di un nuovo percetto, audio-tattile, differente e nuovo rispetto alle componenti tattile e uditiva.

Il problema di Molyneux, reso pubblico da John Locke nel 1694, è stato uno dei primi quesiti circa la natura della rappresentazione mentale e il suo ipotetico ruolo con i processi percettivi. In breve, anche se noto, Molyneux si interroga se un uomo cieco dalla nascita e ora adulto che ha imparato a riconoscere e discriminare un cubo da una sfera tramite il tatto, nel caso in cui riuscisse ad acquistare la vista e avesse posti davanti a lui sia la sfera che il cubo, riuscirebbe a distinguerli guardandoli, ancora prima di toccarli (si veda paragrafo successivo per le evidenze empiriche in risposta al quesito).

Il termine 'rappresentazione' (letteralmente <<qualcosa che sta per qualcos altro>>) è strettamente legato al processo percettivo. Essa di fatti, può essere intesa come una fase successiva al processo di 'filtraggio e organizzazione dell'informazione sensoriale in 'percetto', proprio del processo percettivo. In questo senso, ruolo fondamentale hanno assunto le teorie dell'empirismo e fenomenologiche, che hanno portato a discutere circa il ruolo dei processi bottom-up e processi top-down nella costruzione della percezione e della rappresentazione.

Le teorie che enfatizzano i processi di tipo bottom-up (dal basso verso l'alto) o data-driven (guidate dai dati), pongono l'accento sulla rilevanza che hanno gli stimoli nella costruzione della percezione, ovvero sulle caratteristiche dello stimolo fisico percepito (per arrivare a principi di organizzazione cognitiva dei percetti). Tra le più note accenno la teoria del riconoscimento basato sulla comparazione di sagome; teoria del riconoscimento basato su prototipi; teorie del riconoscimento basato sulla comparazione di caratteristiche; teoria delle descrizioni strutturali (o teoria del riconoscimento per componenti). Per la percezione del parlato le teorie bottom-up vengono dette "teorie passive". Queste teorie non spiegano in modo esauriente gli effetti del contesto.

Le teorie che enfatizzano i processi mentali, semantici o processi top-down (dall'alto verso il basso) sono centrate sui processi cognitivi di ordine superiore (o di alto livello) come principi di organizzazione e concetti, sulle conoscenze preesistenti e sulle aspettative antecedenti che influenzano la percezione, considerando solo in seguito i dati sensoriali (come lo stimolo percettivo). Per la percezione del parlato le teorie top-down vengono dette "teorie attive". Effetto di superiorità delle configurazioni (Pomerants, 1981). Effetto di superiorità degli oggetti (Weisstein e coll., 1982, 1974).

Talvolta i termini 'percezione' e 'rappresentazione' vengono utilizzati come se fossero il medesimo processo. Maurice Merleau-Ponty parla della percezione come di un processo che

ri-crea e ri-costruisce il mondo in ogni momento (Fenomenologia della percezione, p.283), in modo non molto diverso rispetto al ruolo della 'rappresentazione della conoscenza', ovvero del 'formato mentale' che rappresenta la conoscenza di cose, idee ed eventi del mondo esterno (Sternberg, 1996).

In una prospettiva costruttivista-interazionista, con il termine di rappresentazione della realtà si può intendere un processo che riguarda la costruzione della conoscenza o della realtà (Olivetti, 1974). La realtà intesa come costruzione dell'individuo, costruzione mediata da un continuo processo di elaborazione (da cui è stata coniata l'espressione *human information processing*) che comprende sia delle invarianti delle procedure mentali di elaborazione dei dati dell'esperienza, sia elementi varianti e fluttuanti, legati allo stato in cui si trova l'individuo (psichico e fisico, se si considera il rumore e l'oscillazione dell'attività cerebrale). Ne consegue che la rappresentazione mentale può essere intesa, quindi, come un "oggetto mentale dotato di proprietà semantiche" (David Pitt) che può essere considerato da una prospettiva "esterno", se ci si focalizza su ciò che rappresenta, o "interno", se ci si focalizza sullo stato interno dell'individuo che 'lo determina'.

Nel 1943, Kenneth Craik definisce con il 'rappresentarsi' qualcosa che consiste nel tradurre in simboli gli oggetti della percezione e, nel caso, produrre altri simboli, che 'rappresentino' qualche altra cosa. Inoltre, alcune di queste rappresentazioni potrebbero essere tradotte in senso inverso. Infatti, secondo Craik le attività compiute da certi sistemi fisici (in grado di rappresentarsi qualcosa) non sarebbero altro che il risultato di una ulteriore traduzione: quella che, appunto, trasforma i simboli (le rappresentazioni) in azioni (eventi o processi nel mondo). Ne consegue quindi che le rappresentazioni, nella visione di Craik, non sono altro che il medium simbolico che connette certi eventi o processi fisici (l'oggetto della percezione) con certi altri eventi o processi di tipo fisico (le azioni compiute dai soggetti che sono in grado di rappresentarsi qualcosa). Perché un sistema sia in grado di costruire rappresentazioni (nel senso di Craik) sembrerebbe necessario che:

- il sistema sia in grado di rispondere a certe condizioni di tipo fisico, legate ai fatti da rappresentare;
- il sistema sia in grado di trasformare in simboli la suddetta 'risposta'.

Si definiscono quindi, due grandi aspetti della rappresentazione, quello strettamente legato al processo percettivo che include l'aspetto del movimento e si interazione con l'ambiente, e quello prettamente astratto, simbolico, che include l'aspetto comunicativo con l'ambiente. Riprendendo l'espressione di 'rappresentazione della conoscenza', è possibile distinguere due tipi di rappresentazione, legati o alla conoscenza dichiarativa (fatti che possono essere

affermati) o alla conoscenza procedurale (procedure che possono essere messe in atto) (Sternberg, 1996), nonché la distinzione tra il 'cosa' e il 'come' (Ryle, 1949).

Si tenga presente che tutte le forme di rappresentazione non presentano non tutte le caratteristiche della cosa che è rappresentata, ma ciascun tipo di rappresentazione presenta solo determinate caratteristiche distintive. Detto ciò, se si considera la rappresentazione della conoscenza dichiarativa, in letteratura vengono riportate principalmente la distinzione delle rappresentazioni analogiche, e simboliche. Le rappresentazioni analogiche presentano attributi concreti, che riprendono analogicamente caratteristiche e proprietà degli oggetti del mondo reale che rappresentano. Alcune rappresentazioni di questo tipo possono essere esperite sia simultaneamente, che analizzate sequenzialmente (per esempio, zoomare un'immagine mentale). Le rappresentazioni simboliche riguardano la relazione tra una parola e la cosa che arbitrariamente rappresenta, ovvero tra l'oggetto reale che rappresenta e l'etichetta verbale è presente una relazione simbolica, conseguente ad applicazione di una regola.

Tra le teorie più influenti sui due tipi di rappresentazioni della conoscenza dichiarativa, analogica e simbolica, si ricorda la teoria del doppio codice (o *dual-code hypothesis*) di Allan Paivio (1969, 1971), in cui viene estremizzata l'esistenza di informazione di immagini non verbali analogiche, e il linguaggio come informazione verbale simbolica. Diverse evidenze sperimentali supportano questa teoria, tra cui la differente capacità interindividuale nel creare immagini verbali e l'interferenza dei compiti che richiedono abilità immaginative vs. verbali. Si tenga presente che implicitamente, Paivio abbraccia una teoria di equivalenza funzionale tra le immagini mentali e i percetti rappresentati (Farah, 1988; Finke, 1989)

Contestualmente, si è posta in antitesi, la teoria dell'ipotesi proposizionale (o concettuale-proposizionale), sostenuta in primis da Zenon Pylyshyn (REF 1973, 1981, 2003) e Fodor (??? 1975). L'autore sostiene che è possibile ipotizzare l'esistenza delle sole proposizioni, usate per descrivere ogni tipo di relazione (varie proposizioni possono essere combinate per rappresentare relazioni più complesse). Secondo questo autore, le immagini non possono essere analoghe ad eventi fisici, come di fatto non ci sono rappresentazioni fisiologiche di immagini nel cervello, per cui esse sono solo degli epifenomeni (fenomeni secondari che si verificano come risultato di altri processi cognitivi). Anche le immagini, secondo Pylyshyn, utilizzano informazioni simboliche, astratte, immagazzinate in un formato proposizionale (Sternberg, 1996).

Una posizione intermedia è stata proposta da Johnson-Laird (1983, 1989) nella sua ipotesi dei modelli mentali, ovvero secondo l'autore la rappresentazione mentale può avere forme

differenti, includendo quindi sia proposizioni, che immagini, che modelli mentali. In questo modello le proposizioni sono rappresentazioni completamente astratte ed esprimibili mentalmente; le immagini sono rappresentazioni specifiche di oggetti particolari (preservano caratteristiche quali dettagli da un'angolazione particolare); e i modelli mentali sono anch'essi rappresentazioni di tipo analogico di concetti e oggetti, spazialmente e temporalmente analoghi, ma 'visualizzabili' da qualsiasi prospettiva (il percetto in questione viene definito prototipo).

Similmente, anche Wilma Bucci (1997) ipotizza, nella sua teoria del codice multiplo, l'esistenza di almeno tre codici: verbale, non verbale-simbolico e non verbale-non simbolico (o sub-simbolico o connessionista). In questo modello, il codice verbale ovviamente si basa sulle parole o sul linguaggio parlato; il codice non verbale-simbolico utilizza ad esempio immagini, o emozioni, cioè simboli di strutture sottostanti che però non sono espressi in parole; ed il codice non verbale-non simbolico non è nè verbale nè simbolico, quindi non è rappresentabile in nessun modo,e si basa sull'elaborazione distribuita parallela, funziona in modo continuo o analogico, e comunica informazioni che riguardano ad esempio il movimento (come per la conoscenza procedurale), ma anche buona parte del funzionamento interpersonale. L'autrice sostiene che i codici cognitivi multipli sono dei modi con cui diverse parti della nostra mente funzionano, più o meno collegate tra loro con riferimenti reciproci, cioè con quella che la Bucci chiama Attività Referenziale (*Referential Activity*), , con la quale i diversi sistemi o codici riescono a immagazzinare nella memoria a lungo termine. Questo può implicare anche la formazione di nuove connessioni non esistenti precedentemente o andate perdute nel corso del tempo e a causa di difese o esperienze traumatiche per lo sviluppo. In sintesi, questi tre codici sono dei prototipi che si distribuiscono in un *continuum* di modalità di immagazzinare e trasmettere le informazioni della memoria a lungo termine.

Finora si è parlato della rappresentazione della conoscenza dichiarativa. Per quanto riguarda le rappresentazioni della conoscenza procedurale (conoscenza di come organizzare sequenze in passi procedurali per eseguire azioni), alcuni dei principali modelli nascono in ambito dell'intelligenza artificiale, nonché dalle ricerche basate su simulazioni computerizzate e dalla Teoria dell'elaborazione dell'informazione, in cui viene fatta un'analogia mente-macchina (l'uomo non è molto differente rispetto ad un computer). L'interesse per la rappresentazione delle conoscenze procedurali nasce in ambito dell'intelligenza, per scopi principalmente applicativi, ovvero per poter programmare il computer ad eseguire compiti in modo intelligente, in modo che potessero essere ricavati dei logaritmi che consentissero in

modo accurato ed efficiente di manipolare le conoscenze per svolgere dei compiti cognitivi. Sono stati quindi approfonditi i termini di 'concetto', l'unità fondamentale della conoscenza simbolica (idea relativa a qualcosa), e di 'schema', ovvero l'organizzazione mentale comprendente un insieme di concetti interrelati in un'organizzazione significativa. Viene anche coniato il termine di 'rete semantica', per indicare che i concetti o 'nodi' sono collegati tra loro in una rete di connessioni interconnesse da relazioni significative (che possono esere di appartenenza categoriale, attributi o significati semantici).

Una sorta di congiunzione, che unisce in un unico modello le rappresentazioni di conoscenze dichiarative e procedurali (ed in seguito ha anche incluso la memoria di lavoro, intesa come conoscenza attivata disponibile per l'elaborazione cognitiva a capacità limitata) è stato formulato da John Anderson (1976), conosciuto come 'modello ACT' o 'controllo adattivo del pensiero' (*adaptive control of thought*). Il modello racchiude la conoscenza dichiarativa con le 'reti proposizionali' (intendendo con 'proposizione' la più più piccola unità di conoscenza che può esistere come asserzione a sé stante) e la conoscenza procedurale con i 'sistemi di produzione' (dove le procedure sono un 'sapere che' viene trasformato i procedura di implementazione e quindi un 'sapere come').

Per completezza, si ricordi che oltre alle teorie dell'elaborazione dell'informazione ispirate dal computer (per le quali l'informazione viene elaborata in modo seriale), sono nate anche delle teorie e modelli di elaborazione distribuita in parallelo o modelli connessionisti (si veda McClelland e Rumlehart, 1986 e il modello dell'attivazione interattiva, McClelland e Rumenlhart, 1981).

Formato della rappresentazione mentale: modale vs. amodale

Ad oggi, il dibattito sul formato della rappresentazione mentale ha assunto due posizioni: una riguarda la rappresentazione come strettamente legata alla percezione sensoriale, modello modale, ovvero strettamente legata ai sensi che veicolano un certo tipo di informazione iniziale; ed un concetto di rappresentazione più astratto, o legato ad aspetti linguistici o che enfatizzano equivalenze funzionali delle informazioni in entrata, modello amodale o sovramodale. In sostanza, anche nonostante i recenti risultati delle neuroscienze e le nuove conoscenze sul funzionamento neuronale, il dibattito rimane ancora aperto.

Il concetto di modalità è per certi versi confrontabile e similare al concetto di modularità. Ogni modalità può essere considerata, di fatti, come una singola unità di ricezione

ed elaborazione dell'informazione in entrata (proprietà dello stimolo specifiche per modalità). La concezione della mente come modulare la specificità per dominio, è stata formulata da Jerry Fodor, nella sua teoria della modularità (Fodor, 1983 *The Modularity of Mind*). Secondo Fodor, la mente è divisa in moduli separati che operano in modo quasi del tutto indipendente e ciascun modulo può elaborare solo un tipo specifico di input. Secondo l'autore dei trasduttori ricevono lo stimolo iniziale di input e lo trasformano in un formato manipolabile dal sistema cognitivo, il quale viene decodificato dai moduli (o sistemi di input) e trasformato in una rappresentazione utilizzabile dai sistemi centrali deputati alle funzioni superiori (es. Pianificazione del comportamento).

Come citato, se il formato della rappresentazione mentale viene legato come frutto esclusivo di ciascuna modalità, allora si parlerà di modello modale, nel caso in cui si ipotizza una equivalenza funzionale del formato della rappresentazione mentale risultate da qualsiasi modalità sensoriale, allora si parlerà di un modello amodale o sovramodale (si ricordi che modelli di rappresentazione mentale astratta basata sul significato e sui simboli, quali quelli linguistici, supportano questo secondo modello).

In favore del modello sovramodale, il già citato, Craik pensava che tra le rappresentazioni e i fatti rappresentati dovesse sussistere qualche tipo di equivalenza strutturale, ovvero non riteneva indispensabile che una rappresentazione (o un 'modello') somigliasse al suo oggetto; piuttosto, pretendeva che il modello funzionasse in modo tale da produrre l'esito previsto, «per gli aspetti che si stanno considerando in un dato momento».

Per rispondere al quesito di Molyneux già citato, appoggiando il modello sovramodale, si può ipotizzare che il cieco che riacquista la vista possa sperimentare quantomeno una forma peculiare di visione, che riflette le proprietà spazio-temporali comuni, o che sperimenti solamente le 'irritazioni' iniziali del processo visivo (corrispondente alla stimolazione periferica recettoriale) (Giannoli, 2005). In effetti, le evidenze dei pazienti ciechi dalla nascita, operati con 'retine artificiali', hanno riportato di riuscire trasferire l'oggetto della percezione, non percependolo più sulla loro pelle, ma sulla scena posta di fronte a loro (ripresa da una telecamera ordinaria) e di sperimentare il fenomeno dell'occlusione visiva, in funzione della posizione assunta, ma presentano una serie di deficit, tra cui difficoltà nel riconoscere oggetti interi (visto che la scansione alla quale i pazienti erano abituati, quella tattile è di tipo locale), difficoltà nella percezione delle distanze, ombre e delle prospettive, dimenticano facilmente ciò che hanno riconosciuto (inizialmente si aiutano con il tatto), non sono sensibili ad alcune illusioni ottiche e non sperimentano il fenomeno dell'inversione gestaltica, hanno difficoltà ad attribuire significato agli schizzi, momenti simili ad agnosie (risultati riportati da Ackroyd et

al., 1974; Banissoni et al. 1967; 1968; Gregory e Wallace, 1974; MacLeod, 2001; Sacks, 1995; Smallmanet al, 2000; Umezu et al. 1975).

Uno dei primi studi che ha portato evidenze empiriche in risposta al quesito Molineaux, è lo studio svolto da Banissoni e coll. nel 1967 e nel 1968 in cui mostrano che adulti ciechi congeniti dalla nascita, dopo un'operazione in età adulta di osteo-odonto-cheratoprotesi di Strampelli, non sono riusciti a riconoscere le forme geometriche a cui erano stati precedentemente sottoposti con il tatto, con l'utilizzo della vista. Di fatto, anche dopo diverse prove di generalizzazione, i soggetti hanno mantenuto alcuni dei problemi sopra citati.

Di contro, le evidenze citate sulla sostituzione sensoriale, sembrano favorire l'ipotesi di un formato della rappresentazione sovramodale. Si ricordino nuovamente gli studi condotti con risonanza magnetica funzionale, che mostrano una effettiva riorganizzazione delle aree visive, in alcuni soggetti non vedenti. Si noti che, sottoponendo individui ciechi a opportuni stimoli tattili, il sistema visivo risulta attivato, sia a livello della corteccia primaria, che al livello di quella extra-striata, e ai livelli superiori; anche se un'analoga attivazione non si registra invece nei vedenti, sottoposti a un identico stimolo (Giudice et al, 2002).

A livello neurale si può fare una distinzione tra attivazione multimodale e sopramodale. L'elaborazione sopramodale o attivazione cerebrale sopramodale sembra essere causata dalla stimolazione di una modalità sensoriale o dalla stimolazione simultanea di più modalità sensoriali (Klemen e Chambers, 2012). A livello neurale si parla quindi di una necessaria contingenza sia spaziale che temporale degli stimoli delle diverse modalità sensoriali.

La presenza di studi a supporto dell'attivazione di neuroni sopramodali potrebbe dipendere: 1) dall'attivazione sia di una modalità A che attiva i neuroni attivati generalmente dalla modalità B (per la maggior parte dei casi si parla della prossimità spaziale di queste due aree, 2) sia il fatto che l'attivazione delle aree primarie di A co-attiva anche le aree secondarie della modalità B, e viceversa. Inoltre, non bisogna escludere la possibilità che neuroni selettivamente specifici per elaborare la modalità A siano mescolati a neuroni selettivamente specifici per elaborare la modalità B. In passato sono stati definiti neuroni multisensoriali, in neuroni riscontrati nel collicolo superiore dal momento che 1) rispondevano alla stimolazione di diverse modalità sensoriali, 2) mostravano il fenomeno detto della 'superadditività' (risposta maggiore rispetto alla semplice somma delle risposte unisensoriali), 3) mostravano anche il fenomeno dell'effettività inversa (la risposta minima dello stimolo unisensoriale meno effettivo produce una risposta multisensoriale più effettiva).

Focus sulla rappresentazione spaziale.

Finora si è parlato di rappresentazione mentale della realtà, in generale. In questa sede ci si è voluti soffermare su un tipo di rappresentazione mentale, quella dell'ambiente circostante.
La rappresentazione dello spazio è pervasiva e fondamentale nella nostra vita quotidiana. L'elaborazione dell'ambiente è fondamentale per l'orientamento e l'efficiente spostamento, in tutti i compiti quotidiani (Lynch, 1960). Inoltre, la rappresentazione spaziale influenza ed è influenzata dalla mappatura spaziale che implica la capacità navigazionale, la propriocezione, il controllo motorio, l' identificazione dell' estensione e dei confini tra gli oggetti, così come molti percetti astratti non-spaziali, come relazioni sociali e semantiche (M. W. Peters e B. Drake, 2016).

Secondo le teorie proposizionali, le rappresentazioni spaziali sono tipi di modelli mentali (Johnson-Laird, 1980) o modelli situazionali (Van Dijk & Kintsch, 1983).
Se si considerano le strategie e i processi implicati nella costruzione di una rappresentazione spaziale di un ambiente, esistono due categorie principali, conseguenti a elaborazione spaziale egocentrica o allocentrica (Paillard, 1991; Kosslyn, 1994). L'elaborazione spaziale egocentrica si riferisce ai rapporti delle posizioni degli oggetti con il proprio corpo (prevalenza del corpo e delle coordinate aptiche), mentre l'elaborazione spaziale allocentrica si riferisce ai rapporti tra le posizioni degli oggetti (prevalenza delle coordinate viso-motore). Anche se strategie di elaborazione sono flessibili, ovvero è possibile passare da una strategia all'altra, in letteratura vengono riportate due tipi di rappresentazioni spaziali in base alle strategie di elaborazione spaziale utilizzate: i percorsi e le mappe. I percorsi sono costruiti usando la strategia egocentrica (sequenziale), mentre le mappe vengono costruite con la strategia allocentrica (olistica).
Le mappe cognitive (Tolman, 1948) sono delle rappresentazioni interne dell'ambiente fisico, centrate sulle relazioni spaziali. Consistono in rappresentazioni interne che simulano particolari caratteristiche spaziali dell'ambiente esterno (Rumelhert e Norman, 1988). Esistono tre tipo di conoscenza quando si formano e utilizzano le mappe cognitive: 1) conoscenza dei punti di riferimento (*land-mark knowledge*), relativa all'informazione su particolari caratteristiche in una certa posizione (può essere basata su rappresentazioni immaginative o proposizionali), 2) conoscenza del percorso da seguire (*route-road knowledge*), relativa a itinerari specifici (può essere basata su conoscenze dichiarative o procedurali), 3) conoscenza della mappa d'insieme (*survey knowledge*) relativa alla stima delle distanze tra i punti di riferimento, come in una mappa topografica (può essere basata su rappresentazioni immaginative o proposizionali).

Come per la rappresentazione più in generale, anche per la rappresentazione spaziale il dibattito circa l'elaborazione modalità specifica o sovramodale rimane aperto. In particolare,

ci si è chiesti quale ruolo ha l'esperienza visiva nella rappresentazione spaziale (se la visione sia un fondamentale), e di conseguenza se anche le altre modalità non visive possono trasmettere efficacemente l'informazione spaziale.

Nel contesto della rappresentazione spaziale questi due modelli si sono delineati in 'centrati sulla modalità visiva' e modello amodale.

Il modello a favore di una modalità specifica enfatizza l'importanza dell'esperienza visiva nella costruzione del percetto spaziale e nell'acquisizione di concetti spaziali, portando a sostegno di questa teoria, le evidenze sulle difficoltà di ciechi congeniti nello svolgimento di alcune operazioni di manipolazione spaziale (es. Rotazioni mentali,Ungar e coll., 1995)). Il modello amodale, sottolinea l'equivalenza funzionale delle modalità sensoriali, affermando che esiste un sistema di rappresentazione spaziale che può ricevere informazioni diverse da differenti canali sensoriali. A favore di questo modello sono le evidenze dell'esistenza e creazione di rappresentazione mentale diretta, ovvero mediante mezzi indiretti o simbolici, quali le mappe (es. Richardson, Montello & Hegarty, 1999), diagrammi (Bryant, Lanca, & Tversky, 1995; Bryant & Tversky, 1999) , e descrizioni verbali (ad esempio, Avraamides, 2003, Denis & Zimmer, 1992, Ferguson & Hegarty, 1994, Taylor & Tversky, 1992, Bryant, 1992, De Vega et al.).

L'esperienza visiva nella rappresentazione spaziale è stata a lungo considerata fondamentale nella rappresentazione spaziale dal momento che la visione ha il vantaggio di dare una percezione simultanea dell'ambiente. Infatti, anche se i nostri occhi si muovono costantemente, abbiamo una percezione globale dell'ambiente. Questo ci permette di estrarre l'invarianza spaziale, cioè il fatto che alcune relazioni spaziali restano stabili nel tempo (Foulke, 1982; Millar, 1981a). Ad ogni modo, negli anni, si è visto che anche le altre modalità sensoriali possono trasmettere informazioni spaziali (Thinus-Blanc e Gaunet, 1997).

L'empirismo e lo strutturalismo sostenevano che lo spazio fosse un'idea non visiva, un'idea tattile-cinestesica (composta dei ricordi del tatto e dall'attività muscolare) che le nostre esperienze passate ci hanno insegnato ad associare all'indizio di profondità visivo (Hochberg, 1972).

L'analisi della letteratura mostra che la maggior parte delle differenze tra i ciechi e i vedenti sia quantitativa, piuttosto che qualitativa, e riguarda principalmente il carico di memoria di lavoro (Heller, 1989) o la quantità di informazioni simultanee disponibili per la codifica (Thinus-Blanc e Gaunet, 1997, Blanc e Gaunet (1997) richieste dai vari compiti. Infatti, Millar sottolinea la differenza sostanziale tra la reale abilità o competenza nel

rappresentare lo spazio, e le prestazioni richieste dal compito e dal paradigma specifico utilizzato.

Per quanto riguarda il ruolo dei sensi non visivi sulla rappresentazione spaziale, Millar (1994) enfatizza la nozione di convergenza degli 'inputs', cioè il fatto che nessuna delle modalità sensoriali è necessaria e sufficiente per se per la codifica spaziale. Nel 1951, Worchel afferma che le persone vedenti quando riferiscono un'immagine, riferiscono tutti i tipi di informazioni (tattile, cinestetica), e questo vale anche per le relazioni spaziali tra gli oggetti. Secondo un punto di vista più trasversale, Pick (1974) afferma che le informazioni spaziali aptiche e uditive vengono codificate come rappresentazioni mentali.

Si consideri l'importanza delle informazioni tattili e uditive nella codifica e costruzione delle rappresentazioni spaziali. Sia tatto che udito, infatti, sono strettamente legati alla loro sorgente.

Hausfeld, Power, Gorta e Harris (1982) hanno scoperto che è possibile discriminare oggetti e le loro proprietà spaziali, attraverso i suoni. I suoni, infatti, trasmettono le informazioni spaziali di per sé. Si pensi ad esempio a qualcuno che bussa alla porta. Da questo contesto è possibile trarre informazioni circa l'altezza della persona che bussa (un altro esempio simile è una palla che rimbalza da terra). I suoni possono fornire informazioni sulla direzione e la distanza (ad esempio Middelbrooks e Green, 1991). Inoltre, gli esseri umani ascoltano in tutte le direzioni, i suoni sono distribuiti spazialmente e possono essere utilizzati correttamente per attirare e reindirizzare l'attenzione (Ferlazzo et al., Olivetti et al.), consentendo agli ascoltatori di raccogliere una quantità maggiore di dati (Delogu et al.) e separare più sorgenti (Bregman, 1990), minimizzando anche il carico di lavoro.

Anche il senso del tatto consente di estrarre delle informazioni spaziali. Basti pensare al fatto che gli input tattili e cinestetici ci forniscono informazioni costanti sulla posizione del nostro corpo e sullo spazio peripersonale. Gli studi con giudizi di orientamento della barra hanno mostrato l'importanza dell'orientamento del corpo nei giudizi di orientamento dei bar (Kappers e Koenderink, 1999 e Kappers, 1999, 2002, 2003). Inoltre, recentemente, interfacce aptiche sono state utilizzate per rappresentare immagini spaziali complesse (ad esempio, grafici REF 2015).

Studi precedenti hanno mostrato una buona rappresentazione spaziale delle persone cieche attraverso l'esplorazione audio-tattile. Inoltre, ci sono alcune evidenze dalla ricerca sia teorica che empirica che sembrano convergere nell'idea che una combinazione di suoni e stimolazioni tattili funziona meglio di una singola modalità nella visualizzazione non visiva delle informazioni spaziali (Ramloll et al., 2001). I risultati di questi studi (vedasi anche gli studi del

nostro team, si veda la parte sperimentale), corroborano a livello comportamentale le evidenze a livello neurale di attivazione del sistema visivo che viene attivato anche nei ciechi (per esempio Giudice et al., 2002) e corroborano l'ipotesi che informazioni sensoriali da diverse modalità convergono in una rappresentazione sovramodale dello spazio.

2. CONTRIBUTI SPERIMENTALI

2.1 Integrazione audio-tattile e temporalità.

Per poter parlare di integrazione, come abbiamo visto, gli aspetti spaziali e temporali dell'elaborazione sono importanti. In particolare, l'aspetto temporale è fondamentale affinché si possa verificare una integrazione audio-tattile.

In primo luogo, l'aspetto temporale è necessario per far si che ogni tipo di binding crossmodale abbia luogo. Si ricordi la cosiddetta regola temporale di integrazione multisensoriale ("temporal rule of multisensory integration", esempi Calvert et al., 2004; Stein & Meredith, 1993; Stein & Stanford, 2008) per la quale solo stimoli che seguono un registro temporale vicino (ovvero che originano dallo stesso evento) causano un incremento della risposta (ovvero, una scarica maggiore rispetto alla scarica che si verifica per una singola modalità). Gli stimoli

multisensoriali non devono necessariamente essere perfettamente sincroni, ma devono accadere all'interno della stessa finestra temporale di integrazione ("temporal window of integration", si veda Meredith et al., 1987). Diverse modalità sensoriali, infatti, presentano delle peculiarità in termini di velocità di conduzione, latenza delle risposte e tempo di elaborazione neurale, ragion per cui l'integrazione delle informazioni delle diverse modalità deve riuscire a superare queste differenze specifiche (Lestienne, 2001; Nicolas, 1997; Vroomen & Keetels, 2010). L'esistenza di una certa 'tolleranza' temporale ha non solo implicazioni empiriche ma anche teoriche (per esempio, si veda la teoria dell'elaborazione cognitiva di Pöppel, 2009; van Wassenhove, 2009).

In secondo luogo, l'aspetto temporale (in termini di successione di eventi) è uno degli elementi che accomuna queste due modalità sensoriali. Si ricordi, infatti, che sin dall'inizio si è detto che entrambe le modalità codificano ed elaborano le informazioni dell'ambiente circostante in modo sequenziale; inoltre l'esplorazione tattile e l'ascolto uditivo implicano una scoperta sequenziale dell'ambiente circostante; infine, entrambe presentano un aspetto 'cumulativo' degli eventi. Si considerino gli studi sui 'treni di stimoli' (si veda Ocelli et al., 2011).

Fatta questa premessa, la ricerca bibliografica effettuata in merito agli studi sull'integrazione audio-tattile ha portato a verificare l'assenza di studi che indagano la finestra di integrazione audio-tattile in flussi continui audio-tattili. Si è deciso quindi, di studiare il fenomeno dell'integrazione audio-tattile in flussi ritmici continui in cui fosse manipolata una piccola finestra temporale dove i suoni sono asincroni. Inoltre, si è pensato che questo doveva essere un processo 'percettivo', automatico (quindi relativo a finestre temporali molto piccole) se avveniva entro la finestra di integrazione temporale, a cui si è fatto cenno. Si è voluto creare quindi un metodo per confrontare questo appaiamento puramente percettivo con uno di tipo cognitivo. Per far ciò, si è pensato di 'distanziare' i flussi l'uno dall'altro, chiedendo ai partecipanti di continuare a 'pensare' ai due flussi come sincroni (come un flusso, ma con stimoli di un flusso differiti) e di individuare le asincronie. In questo modo ci si aspetta di distinguere quando il compito è di tipo 'percettivo' rispetto a quando è di tipo 'cognitivo'(il che in termini di performance si riscontra in una maggiore individuazione delle asincronie). Inoltre, si è deciso di svolgere anche due condizioni unisensoriali, per mettere a confronto la performance con la finestra temporale audio-tattile. Ci si aspetta che le finestre temporali entro una singola modalità (uditiva e tattile) siano più piccole rispetto alla finestra temporale di integrazione di due modalità, in questo caso, l'integrazione audio-tattile (si vedano i dati già riportati nella parte teorica, per binding uditivo, REF e binding tattile REF!). Ciò implica che usando le stesse finestre temporali per gli studi sia uni che pluri sensoriali, la performance

unisensoriale per le finestre temporali maggiori dovrebbe risultare notevolmente peggiore rispetto a quella multisensoriale.

2.1.1 Studio sulle finestre temporali: detezione di incongruenze temporali in flussi ritmici.

Come abbiamo visto nella parte introduttiva, la temporalità nei processi di binding è fondamentale affinché il nostro organismo percepisca due eventi come sincronizzati o meno. Si è fatto cenno circa il fatto che ogni modalità sensoriale ha una finestra temporale specifica entro cui gli stimoli vengono percepiti in sincrono o distinti. Si è visto, inoltre, che l'integrazione crossmodale presenta anche delle peculiarità in merito alle finestre tattili. In particolare, dal momento che l'integrazione audio-tattile si verifica di continuo nella vita quotidiana (per esempio sbattere su un tavolo e sentirne sia suono che vibrazione), si può assumere che vi sia un'associazione naturale, o intrinseca, che avviene tra queste due modalità, in modo puramente percettivo. Se quindi, queste due modalità si verificano oltre la finestra temporale, o in altro modo, viene inserita una 'distanza' tra l'evento uditivo e tattile (o viceversa), i due eventi risulteranno come distinti da un punto di vista percettivo, ma potranno essere confrontati con uno sforzo cognitivo.

In questo studio, ci si è chiesti se quando si presentano due flussi ritmici continui sincronizzati per la maggior parte del tempo, i soggetti sono in grado di individuare delle piccole incongruenze temporali tra i due flussi (detezione di tipo percettivo), anche quando questi sono presentati 'a distanza' nel tempo (detezione di tipo mnestico-immaginativo), in tre condizioni: due in cui i due flussi appartengono alla stessa modalità (condizione uditiva e condizione tattile) e una in cui i due flussi appartengono a modalità differenti (condizione audio-tattile). L'aggiunta di una 'distanza' temporale tra i flussi è stata inserita per confrontare la prestazione tra la situazione in cui i due flussi sono realmente sincronizzati per la maggior parte del tempo, per cui i partecipanti rilevano le incongruenze ad un livello percettivo (si parla di incongruenze di 100 ms, 250 ms e 400 ms), rispetto ad una condizione in cui ai partecipanti è chiesto di 'sforzarsi di pensare ai due flussi come sincroni', e quindi di rilevare le piccole incongruenze facendo fronte ad un carico di tipo cognitivo. Assumendo che la detezione delle asincronie sia automatica e presenti una finestra temporale molto piccola, ci si aspetta che nel compito in cui i due flussi uditivi sono presentati in simultanea, la performance di accuratezza sia molto alta. Inoltre, per lo stesso motivo, ci si aspetta una performance peggiore (rispetto a quella ottenuta nello studio precedente) nelle condizioni differite.

Dal momento che esistono tre livelli di asincronia (a 100 ms, 250 ms e 400 ms), si suppone che man mano che i due flussi sonori vengono distanziati, sia più difficile per i partecipanti rilevare le asincronie piccole.

Le due condizioni unisensoriali e multisensoriali, sono rilevanti ai fini di un confronto tra le finestre temporali specifiche per ciascuna modalità (uditiva e tattile) e la finestra temporale crossmodale, che si suppone possa comportare ad una prestazione intermedia. Infatti, qualora sia presente una differenza tra le due modalità sensoriali, si potrà ottenere una prestazione intermedia. Non si può escludere la possibilità di una prestazione migliore, nel caso il cervello 'segua' la modalità sensoriale più alla portata per rispondere, ovvero non si possono escludere gli effetti di interferenza o di facilitazione nell'utilizzo di entrambe le modalità nel confronto tra i due flussi ritmici. Si ipotizza che la finestra temporale di integrazione dei due flussi unimodali sia più piccola rispetto all'integrazione multimodale audio-tattile, e che la finestra temporale tattile sia più grande rispetto a quella uditiva (REF). Questo implica che nella modalità tattile (rispetto a quella uditiva) si verifichi maggiore difficoltà a rilevare le incongruenze più piccole.

Per ogni condizione, sono stati svolti dei piloti con 3-4 soggetti, per in cui sono stati manipolati i livelli di asincronie tra i due flussi (sono stati provati 4 livelli ripetuti 10 volte ciascuno: 100 ms, 200 ms, 300 ms, 400 ms). Dal momento che ogni asincronia deve essere presente almeno 10-15 volte per poter dare una stima dell'effettiva detezione dell'asincronia (REF negli studi di psicofisica) e data la poca differenza tra le finestre temporali scelte, si è portato a tre il numero delle asincronie (a 100 ms, 250 ms e 400 ms). In questo modo si è cercato di favorire un quadro ampio, e una durata delle sessioni non troppo lunga. Durante i piloti (2 soggetti ciascuno) sono stati delineati dettagli circa la struttura di ogni sessione (si è reputato utile introdurre qualche minuto di adattamento e mostrare degli esempi ai soggetti).

Per la stimolazione uditiva, è stata tarata, nei pilota, la durata dei beep, dal momento che a 'beep' eccessivamente lunghi, il compito risultava molto facile anche a finestre temporali molto piccole. La spiegazione che ci si è dati è che stimoli più lunghi erano più salienti (REF).

Per la stimolazione tattile è stata provata, nei pilota, le stimolazione dei due indici (destro e sinistro), del dito indice e anulare della mano destra e dei due pollici (destro e sinistro). Si è visto che la stimolazione di dita appartenenti a mani differenti fosse più semplice rispetto alla stimolazione di dita appartenenti alla stessa mano (REF). Si è scelta quindi, la stimolazione dei due pollici, sia per chiarezza dello stimolo che per praticità nel maneggiare il dispositivo, che viene tenuto come un joystick; mentre le risposte vengono date premendo (la barra spaziatrice con l'anulare destro.

Metodo.

Soggetti.

Hanno preso parte all'esperimento 30 partecipanti (femmine, maschi, età media), 10 di essi hanno svolto la condizione tattile, 10 la condizione uditiva e 10 quella audio-tattile. Tutti i partecipanti hanno dichiarato di essere in buono stato di salute, di non aver deficit uditivi o di sensibilità tattile delle dita, e di non aver avuto precedenti diagnosi psichiatriche. I partecipanti hanno preso parte all'esperimento volontariamente. Per ogni condizione, metà dei soggetti ha iniziato con la sessione con i flussi ritmici sincronizzati (distanza 0ms) per poi svolgere quelli a distanza di 150 ms e 300 ms; e metà dei soggetti ha iniziato con i blocchi in ordine inverso (blocco con gli stimoli distanti 300 ms e poi distanza di 150 ms e sincroni).

Strumenti e apparati

Sono stati utilizzati un comune computer portatile, cuffie "Marshall Monitor" e lo strumento che abbiamo chiamato '*Feelspeech prosthesis*'. Per la condizione audio-tattile, è stato utilizzato il *Feelspeech* con i pin a matrice di 4x4 (pollice destro) (si veda fig. 4 b) in appendice), mentre per la condizione solo tattile è stata usata la matrice 2x4 per ciascun pollice (si veda fig. 4 a) in appendice). Il dispositivo viene tenuto con l'intera mano (come un joystick) e il palmo dei pollici viene posto sulla matrice.

Stimoli.

<u>Stimoli uditivi.</u>

Nella condizione audio-tattile, come stimolo uditivo viene utilizzato il rumore rosa, la cui presenza-assenza porta a percepire un flusso ritmico incostante. In altre parole, vengono utilizzate (in modo random) quattro pause (da 100, 200, 300, 400 ms) tra uno stimolo e l'altro; e ciò crea a livello percettivo un suono continuo simile ad un codice morse.

Nella condizione uditiva, viene presentato sia il rumore di fondo (che i soggetti devono ignorare) che viene creato prendendo il rumore bianco e filtrandolo passa-banda a 1000 e 1200 Hz (1100 Hz è la media geometrica di 200 Hz e 6100 Hz), sia due beep (di 200 Hz e 6100 Hz che seguono un ritmo random. Ciascun beep dura 30 ms e vi sono quattro tipi di pause tra un suono e l'altro (che creano percettivamente il senso del ritmo che cambia, come per la condizione audio-tattile). Si è scelto di utilizzare un rumore di fondo, dal momento che studi sull'acustica (conversazione con Daniel Pressnitzer) mostrano che quando si presentano due suoni in cui è presente molta differenza di 'contenuto' di frequenze (ovvero un suono molto

acuto e uno molto grave) vi sono una serie di armoniche intermedie, per cui si può avere un effetto percettivo di un suono intermedio tra i due. Il rumore di fondo, quindi, consente di mascherare ogni possibile 'falso' segnale creato dalla combinazione delle armoniche dei due suoni (ed evitare, di conseguenza, che i partecipanti decidano se i suoni sono asincroni o meno in base all'ascolto della presenza/assenza del suono intermedio che si verrebbe a creare percettivamente).

Stimoli tattili. Per la stimolazione tattile è stato utilizzato il dispositivo *FeelSpeech*.

Per entrambe le condizioni, solo tattile e audio-tattile, i partecipanti tengono il dispositivo nelle mani, come un joystick, con la differenza che nella condizione audio-tattile, la stimolazione tattile viene effettuata solo sul pollice destro (e come già citato, viene utilizzata una matrice di stimolazione con 4x4 puntine), mentre nella condizione solo tattile, la stimolazione è presentata ad entrambi pollici con una matrice di stimolazione 2x2.

Nella condizione audio-tattile viene veicolata l'informazione che corrisponde all'ampiezza generale dell'informazione uditiva (vengono distinte 4 range di frequenze a 0-650 Hz, 650-1400 Hz, 1400-3000 Hz, 3000-5000 Hz, 5000-6500 Hz). In pratica, ad ampiezza acustica maggiore, viene associato un maggior numero di 'puntine', ovvero la sensazione di percepire più 'spilli' sul palmo del pollice. Quindi, ad un basso livello di ampiezza, si alzano su solamente le prime due file di array di 'puntine' o 'bottoni', ad un livello medio di ampiezza, metà file della matrice, e ad ampiezza massima, tutte le 'puntine' sono alzate. A livello visivo, se si guarda lo strumento, è un pò come vedere le barre che nei vecchi stereo indicavano il volume e l'ampiezza dei bassi (si veda fig. 13 dell'appendice).

Nella condizione solo tattile, tutti i puntini, per ciascun pollice (matrice 2x4) si alzano e si abbassano e i soggetti confrontano la stimolazione dei due pollici (puntini alzati o abbassati in sincro).

In tutte le condizioni, come accennato, gli stimoli sono sincronizzati per la maggior parte del tempo, sia che i flussi vengano presentati simultaneamente, sia quando vengono distanziati (di 150 ms e di 300ms) e i soggetti devono rilevare la presenza saltuaria di tre diversi livelli di asincronia (15 prove per ciascuna, per un totale di 45 asincronicità): a 100 ms, 250 ms e 400 ms.

Procedura.

Come già detto i partecipanti sono stati divisi a gruppi di 10 (in modo casuale) per svolgere un tipo di condizione: solo tattile, solo uditiva, audio-tattile.

Indipendentemente dal tipo di condizione, ogni partecipante svolge tre blocchi sperimentali di 13 minuti ciascuno, con 5-10 minuti di pausa tra una sessione e l'altra. La durata complessiva è di un'ora a partecipante. Ogni blocco è composto da un minuto di adattamento, poi vengono presentati 4 esempi di incongruenza (gli esempi sono indicati da un 'BIP' breve acuto, 3 s prima di ciascun esempio), e dopo viene presentato un 'BIP' grave e lungo che indica l'inizio del test vero e proprio che dura 12 minuti. Metà dei partecipanti (per ciascun gruppo) inizia con la condizione in cui i due flussi sono presentati simultaneamente (e poi svolge il blocco con i flussi distanziati di 150 ms e poi di 300 ms), e metà inizia con i blocchi in ordine inverso per evitare che i risultati vengano influenzati dalla stanchezza. Il compito dei partecipanti, per tutti i blocchi, è quello di premere la barra spaziatrice ogni qualvolta avvertono che è presente un'asincronia tra i due flussi.

Nei blocchi in cui gli stimoli sono presentati simultaneamente, si suppone che le incongruenze vengano immediatamente avvertite, mentre nei blocchi in cui i due stimoli vengono presentati ad una distanza, costante lungo tutta la sessione, si presume una difficoltà maggiore nel confrontare la temporalità dei due flussi (si veda fig. 16 dell'appendice).

Risultati

Per separare la distribuzione del segnale dal rumore (la presenza di molti falsi allarmi), è stao analizzato il il valore *d'* o *dee-prime*.

Per quanto riguarda la condizione in cui i due flussi sono presentati in sincrono (distanza 0 ms), i soggetti hanno individuato le incongruenze per tutti i livelli di asincronia (valore d'>1.5) solo per la condizione solo uditiva (rispettivamente, valori d'=3.23; 3.76, 3.43); mentre nella condizione solo tattile e audio-tattile sono riusciti a rilevare solo per asincronie di 250 ms e 400 ms (rispettivamente, valori d' della condizione solo tattile 1.73; 2.45 e valori d'della condizione audio-tattile 2.32, 2.21).

Per riguarda la condizione in cui i due flussi sono presentati distanziati di 150 ms e di 300 ms, i soggetti non hanno rilevato le incongruenze a tutti i livelli di asincronia (si vedano fig. 20 e 21).

Discussione e conclusione.

Come detto in precedenza, in questo studio, si è cercato di verificare la capacità di detezione di incongruenze temporali, in stimoli appartenenti alla stessa modalità (uditiva o tattile) vs. stimoli appartenenti a modalità differenti (audio-tattile), in compiti di detezione 'percettiva' o automatica vs. detezione di tipo cognitivo.

I risultati hanno mostrato che quando i due flussi sono presentati simultaneamente, i soggetti riescono a individuare tutti e tre i livelli di incongruenza solo per la condizione solo uditiva, mentre per le condizioni solo tattile e audio-tattile, vengono rilevate le asincronie più grandi (250 ms e 400 ms). Questo dato mostra quindi, che il compito solo uditivo con flussi sincroni, è stato più facile rispetto alle condizioni solo tattile e audio-tattile, confermando che la finestra di elaborazione uditiva (rilevazione di due stimoli uditivi) sia più piccola rispetto a quella tattile e audio-tattile. Una interpretazione che può essere data, quindi, è che dal momento che la modalità uditiva ha una finestra di elaborazione molto piccola rispetto a quella tattile, di conseguenza, la performance audio-tattile è stata influenzata dai tempi di elaborazione della modalità tattile, o in altre parole, la scarsa performance audio-tattile riflette l'influenza dell'elaborazione tattile nei tempi di elaborazione audio-tattile. Questo implica che se riguardo la distinzione tra le finestre temporali di integrazione intra e inter modale, si era ipotizzato che la finestra di integrazione intermodale potesse essere o maggiore rispetto alla finestra di integrazione intramodale, o includerle entrambe, i risultati di questo esperimento sembrano confermare la seconda ipotesi. Infatti, si può ipotizzare che la performance del compito audio-tattile sia 'intermedia' a quella della condizione solo uditiva e solo tattile, dal momento che presenta una finestra temporale 'intermedia' (e di fatto più vicina alla performance della condizione solo tattile), e quindi che la finestra di integrazione audio-tattile si colloca tra la finestra temporale di elaborazione dell'informazione uditiva e dell'informazione tattile.

Inoltre, i risultati mostrano che nei blocchi distanziati di 150 ms e di 300 ms, i soggetti non sono riusciti ad individuare le asincronie. La differenza tra questi due blocchi e il blocco in cui gli stimoli sono presentati in modo sincrono, conferma la distinzione tra un compito di tipo automatico, in cui la detezione delle incongruenze è stata di tipo 'percettivo', con i compiti che hanno richiesto uno sforzo cognitivo, in cui è chiesto un confronto tra le informazioni temporali dei due flussi e l'utilizzo di strategie cognitive, indipendentemente dal tipo di condizione. La maggiore difficoltà nell'individuazione delle asincronie, quindi, viene interpretata come prova circa il fatto che il fenomeno non è puramente percettivo e che i due flussi sono quindi chiaramente oltre la finestra temporale richiesta per percepire in modo automatico le asincronie. Ci si sarebbe aspettata una differenza maggiore tra i due blocchi, ovvero una maggiore difficoltà del compito in cui i flussi erano distanziati di 300 ms, rispetto al blocco di 150 ms, ma l'assenza di tale riscontro, potrebbe fare presumere un 'effetto tetto', ovvero la performance era così scarsa in entrambi i compiti che non è possibile notare una differenza tra i due nella performance registrata.

Una delle possibili critiche che può essere fatta a tal proposito, è l'ovvietà del risultato stesso che può essere interpretata come frutto di un 'incremento di rumore' delle condizioni con distanziamento dei flussi, oppure come 'incremento carico di lavoro' per le stesse condizioni. Si potrebbe infatti obiettare che la buona performance della condizione sincrona rispetto alle altre due, sia dovuta al fatto che non vi sia alcun carico di memoria, mentre nelle altre due condizioni è presente un carico di memoria. In realtà questa critica non è contro la nostra ipotesi di ricerca, per la quale è previsto il fatto che maggiore è il carico cognitivo richiesto per svolgere un compito, più difficile risulta il compito. E ancor di più, la condizione a 'carico di memoria pari a zero' corrobora il concetto di automaticità e il fenomeno di 'puramente percettivo', come da noi formulato.

Conclusione sull'integrazione audio-tattile e temporalità.

Lo studio svolto in questa sede, ha avuto come obiettivo quello di indagare le finestre temporali di flussi ritmici tattili, uditivi e audiotattili, entro cui i soggetti sono in grado di individuare delle incongruenze temporali tra i due flussi. La letteratura esistente, infatti, non mostrava alcuna evidenza sperimentale circa i tempi di detezione di incongruenze temporali tra flussi con ritmi non costanti, nelle tre condizioni indicate. Inoltre, sono stati presentati un compito definito 'percettivo', ovvero in cui la detezione delle incongruenze temporali avviene in modo immediato; ed due condizioni 'cognitive', in cui per individuare le incongruenze temporali, i soggetti devono utilizzare strategie cognitive.

Come visto, i risultati hanno confermato la letteratura precedente circa la velocità di elaborazione uditiva, che risulta essere più veloce rispetto a quella tattile (REF) anche per la detezione di incongruenze temporali in flussi ritmici. Inoltre, si è riscontrato che la capacità di detezione degli stimoli audio-tattili riflette e si colloca in una posizione intermedia, rispetto alla capacità di detezione delle incongruenze temporali dei flussi ritmici solo uditivi e solo tattili. Questo dato risulta essere molto rilevante ai fini degli studi successivi per il binding audio-tattile. Infatti, si può assumere che una possibile finestra di integrazione di flussi ritmici audiotattili ricada a metà tra la velocità di elaborazione degli stimoli uditivi e quelli tattili. Sono necessari ulteriori studi in merito.

Infine, è stata corroborata la distinzione tra i compiti in cui viene richiesto l'utilizzo di strategie cognitive mnestico-immaginative, per le quali è difficile rilevare perfino incongruenze entro finestre temporali grandi (400 ms), rispetto ai compiti di tipo automatico o 'percettivo', in cui anche la rilevazione di incongruenze molto piccole (100 ms) risulta essere facile.

2.2 Integrazione audio-tattile con stimoli linguistici.

Dalla rassegna della letteratura abbiamo visto che l'integrazione audio-tattile è presente comunemente nella vita quotidiana. Si è visto che probabilmente uno dei motivi per cui stimolazioni uditive e tattili vengono integrate può dipendere dal fatto che essi condividono alcune analogie fisiologiche (es. von Bekesy, 1959), una contingenza delle aree neuronali (es. Ro et al., 2007). Si è visto che esistono molte evidenze in merito all'integrazione audio-tattile in studi in cui si presentano informazioni semplici (una vibrazione o un tono musicale), ma fino ad oggi nessuno degli studi riportati sulla sostituzione sensoriale (studi sulla comprensione 'tattile' del linguaggio per popolazioni di sordi e deficit acustici) ha mai indagato se l'informazione tattile e uditiva di stimoli complessi come il linguaggio, avviene in modo automatico o 'percettivo'. Questi studi infatti, riportano un vantaggio per quanto riguarda le presentazioni multimodali vs. unimodali, non sono una evidenza di una facilitazione né tanto meno dell'integrazione audio-tattile. Al momento quindi, non esistono in letteratura evidenze sulla genuinità del fenomeno di integrazione audio-tattile utilizzando stimoli linguistici.

Di seguito verranno presentati uno studio completo e uno studio pilota in cui si è cercato di trovare evidenze dell'esistenza dell'integrazione audio-tattile a livello puramente percettivo, utilizzando stimoli linguistici semplici, quali le sillabe.

Confermare questa automaticità di integrazione audio-tattile impiegando stimoli linguistici è importante non solo per comprendere i meccanismi di integrazione audio-tattile di stimoli complessi e 'non intrinsecamente' associati (ovvero la loro integrazione richiede un periodo di apprendimento, visto che non si tratta di caratteristiche la cui integrazione non è presente in natura e di cui non si ha mai avuto esperienza in precedenza), ma è importante anche per le applicazioni pratiche rilevanti per lo sviluppo di sistemi di sostituzione sensoriale tattile per aiutare popolazioni speciali (non udenti o popolazioni con deficit uditivi) a comprendere il linguaggio.

2.2.1 Esperimento psicofisico sulla percezione delle sillabe "BA-VA"

Come accennato, nonostante le evidenze in letteratura sull'interazione audio-tattile per domini non-linguistici, la maggior parte degli studi sulla sostituzione sensoriale audio-tattile che cercano di indagare l'influenza (facilitazione) del tatto nella percezione del parlato, hanno trascurato l'importante domanda circa il fatto se questa integrazione audio-tattile effettivamente possa verificarsi con gli stimoli linguistici (in altre parole, ci si dovrebbe interrogare se questa integrazione possa essere di tipo automatico). Di conseguenza, mentre fino ad oggi l'utilizzo della stimolazione tattile era limitato a dimostrare che un particolare dispositivo tattile poteva essere utile per comprendere il discorso, questo richiedeva un lungo sforzo cognitivo e una grande quantità di addestramento. Per verificare se la stimolazione tattile possa influenzare la percezione del linguaggio, e quindi in un certo senso di 'sentire meglio' chi parla, senza implicare un grande sforzo cognitivo, si è deciso di verificare l'influenza della stimolazione tattile sulla percezione uditiva di unità linguistiche estremamente semplici, quali le sillabe. In particolare, sono state scelte le sillabe /ba/ e /va/ dal momento che in letteratura si è visto che queste due matrici tendono ad essere confuse (si veda la matrice di confusione di sillabe di Miller and Nicely, 1955, fig. 3 in appendice). Dal momento che la l'associazione della stimolazione tattile con quella uditiva (sono state utilizzate delle stimolazioni con direzioni differenti associate a sillabe diverse) è puramente arbitraria, sono state inserite delle fasi di training, in modo tale che questa associazione potesse diventare 'automatica'. Per verificare l'influenza del tatto sull'ascolto delle sillabe, alcune sillabe venivano presentate in modo non chiaro, in modo similare ad un inizio di perdita dell'udito (sono state scelte delle sillabe le cui consonanti iniziali vengono confuse da soggetti con deficit uditivi). Inoltre, per verificare che il tatto influenzi l'udito in modo automatico e puramente percettivo, anche quando non viene preso in considerazione consapevolmente o meno, è stata utilizzata una stimolazione tattile sopra e sotto soglia. Infatti, se la stimolazione tattile influenza la percezione

uditiva anche quando le informazioni tattili non vengono elaborate consapevolmente (stimolazione tattile sotto soglia e partecipanti ignorano la stimolazione tattile), si ritiene che questa possa essere una prova di un'interazione audio-tattile a tutti gli effetti. Per verificare questa ipotesi, dopo le varie fasi di training si è testata l'influenza del tatto sulla percezione uditiva con stimolazione tattile sotto soglia (condizione 'test'), e sono state aggiunte due test in cui si è fatto si che il tatto fosse in antitesi con l'informazione uditiva (la direzione associata ad una determinata sillaba, non corrisponde con la sillaba effettivamente presentata), sia con stimolazione tattile sotto soglia (chiamata 'biased weak'), sia con stimolazione tattile sopra soglia (chiamata 'biased strong'). In questo modo, quindi, si è cercato di verificare che la risposta dei soggetti della sillaba percepita fosse influenzata dal tatto, anche quando veniva chiesto di ignorarlo (consapevolmente e a fatica, doveva essere ignorato nella condizione 'biased strong', sia in modo inconsapevole, o sotto soglia, nella condizione 'biased weak').

In questo studio psicofisico, ci si aspetta, quindi, che le informazioni tattili influenzino la discriminazione delle sillabe poco chiare (ovvero che vi sia un cambiamento della soglia) e che questa influenza avvenga in modo automatico, anche quando gli input tattili sono sotto soglia. Inoltre, ci si aspetta che il tatto influenzi la percezione delle sillabe, anche quando i soggetti non prestano attenzione alla stimolazione tattile e perfino anche quando le informazioni tattili sono parzialmente in contrasto con le informazioni uditive. In altre parole, ci si aspetta che la soglia uditiva sia inferiore nella condizione audio-tattile rispetto alla condizione solo uditiva (riflettendo un effetto facilitatorio della modalità tattile). In altre parole, ci si aspetta di ottenere una funzione psicometrica con pendenza maggiore nella condizione audio-tattile rispetto alla funzione psicometrica solo uditiva, confermando la nostra ipotesi circa l'influenza del tatto sulla percezione delle sillabe ascoltate.

Prima di svolgere l'esperimento, sono stati svolti diversi studi pilota per selezionare gli stimoli uditivi (selezione delle sillabe registrate, si è partiti da 21 stimoli, e ne sono stati scelti 13, dallo stimolo 8 allo stimolo 20) e il volume di presentazione degli stimoli uditivi.

Metodo

Soggetti

Hanno preso parte all'esperimento 15 soggetti normoudenti (12 maschi, 3 femmine, età media 36). I soggetti sono di madrelingua diverse, ma tutti usano la lingua inglese nella loro vita quotidiana. Tutti i soggetti hanno firmato la dichiarazione di consenso informato e hanno dichiarato di essere in buona salute e di non avere problemi di udito. I soggetti sono stato pagati 10 euro l'ora (*ERC PoC Grant FeelSpeech*).

Per l'analisi finale, 7 soggetti (soggetti numero: 4, 5, 6, 11, 12, 13, 15) sono stati esclusi, ragion per cui l'analisi statistica ha incluso i restanti 8 soggetti (7 maschi, 1 femmina). I soggetti 4, 5, 11, 12, 13, 15 sono stati esclusi per la loro soglia tattile (valore del sigma troppo elevato, il che significa che la loro percezione tattile non era mai chiara, anche a stimoli tattili la cui velocità era molto lenta) mentre un soggetto (soggetto 6) è stato escluso per la sua soglia uditiva (il valore del sigma è troppo alto, il che significa che presentava difficoltà di discriminazione delle sillabe ai due estremi del continuum, ovvero sillabe molto scandite).

Strumenti

Sono stati utilizzati un comune pc portatile, le stesse cuffie dello studio precedente 'Marshall Monitor' e il dispositivo chiamato *FeelSpeech*. Il dispositivo *Felspeech* presenta solo due stimolazioni tattili, ovvero il movimento (chiamato *microswiping*) di una sola riga di bottoni della matrice 4x4 (riga bassa, si veda FIG) in due direzioni (destra e sinistra). (si vedano fig. 4 b) e c) in appendice).

Stimoli

Stimoli uditivi: abbiamo utilizzato 13 stimoli (come detto, nel pilota erano 21 stimoli, da cui ne sono stati selezionati 13, da stimolo 8 a stimolo 20). Per creare un continuum percettivo da "ba" a "va", con agli estremi le due sillabe udite chiaramente, e nella parte intermedia le sillabe pronunciate in maniera poco chiara. Sono state registrate le sillabe con una voce naturale maschile che pronuncia "ba" e "va". I due file audio sono stati sincronizzati, allineando il punto di partenza della pronuncia della consonante (prima e dopo la sillaba c'era silenzio). Inoltre, per ogni sillaba si è individuato il punto d'inizio dela la vocale / a /. Per creare il continuum, si è scelto di prendere la sillaba più lunga /va/ e manipolare il punto in cui tagliare la traccia audio di /va/ e inserire la traccia audio di /ba/, ottenendo così stimoli ambigui intermedi che iniziano con la sillaba /va/ e si trasformano nella sillaba / ba / (nello stimolo intermedio del continuum vi è una sovrapposizione di 30 ms tra le sillabe). In altre parole, il primo stimolo inizia con la /v/ della sillaba originale / va / e la vocale /a/ presa dalla sillaba originale /ba /, per cui i soggetti percepiscono la sillaba / va / che non è quella originale, ma si sviluppa su due parti delle sillabe registrate originariamente, mentre l'ultimo stimolo presenta tutta la traccia della registrazione del /ba/ iniziale. Inoltre, è stato applicato anche un filtro passo-basso di 1850/1600. Questo filtro elimina le frequenze più alte (che si ricordi che sono le prime frequenze perse dalle persone con problemi di udito).

Gli stimoli sono presentati con volume di 32dB.

Stimoli tattili: sono stati utilizzati 13 stimoli con un movimento (detto da noi *microswiping*) verso sinistra o verso destra. Come accennato, è stato usato il modello di Feelspeech a matrice 4x4 bottoni, ma solo la riga inferiore di 4 bottoni si alzava e abbassava secondo il movimento (si veda fig 2, b) per vedere il movimento). Per creare un continuum di complessità differente degli stimoli tattili, è stata manipolata la velocità del movimento dei bottoni, ovvero la velocità di innalzamento e abbassamento di ciascun bottone. Ogni bottone rimane alzato per 20 ms prima di scendere (rientrare nel dispositivo). Si è variato, quindi, il divario tra il tempo in cui ciascun bottone si alza e il bottone precedente si abbassa (ovviamente si deve considerare la direzione del movimento, per cui il primo bottone sarà quello in basso a destra o sinistra nell'ultima fila del dispositivo). In sintesi, quindi, i perni possono alzarsi in modo indipendente, parzialmente sovrapposti, o in sincro. Per esempio, a tempo 0 ms, i pin sono tutti alzati. Il tempo tra un pin e l'altro può assumere i seguenti valori: -40ms, -30ms, -20ms, -15ms, -10ms, -5ms, 0ms, 5ms, 10ms, 15ms, 20ms, 30ms, 40ms (valori negativi indicano il movimento verso sinistra, mentre i valori positivi indicano il movimento verso destra). (si veda fig. 5 dell'appendice)

Stimoli audio-tattili: sono stati utilizzati tutti e 13 gli stimoli uditivi, associati di volta in volta a solo due stimoli tattili, che cambiano in ogni sessione di allenamento e di test, in base alla soglia (sopra e sotto soglia, si veda la procedura).

Procedura

L'esperimento prevede per ogni soggetto: misura della soglia uditiva, misura della soglia tattile, varie sessioni di addestramento (9 blocchi), tre condizioni di test e nuovamente la misura della soglia uditiva. L'intero esperimento ha durata complessiva di un'ora e mezza a soggetto (il tempo è stimato escluse le pause). Dopo ciascun blocco è stato chiesto ai soggetti di riportare se vi fossero eventuali cambiamenti nella percezione uditiva (si veda fig. 23).

La misura della soglia uditiva include un blocco di familiarizzazione e un blocco principale. Nel blocco di familiarizzazione, i soggetti ascoltano tre stimoli (numero degli stimoli: 1, 11, 21, dove lo stimolo numero 1 indica un chiaro / va / e lo stimolo numero 21 è un chiaro / ba /), ciascuno ripetuto tre volte (in ordine casuale). I soggetti devono rispondere se hanno ascoltato la sillaba / ba / o / va / premendo il pulsante '1' per / ba / e il pulsante '2' per / va / (pulsanti di una comune tastiera di pc portatile). Nel blocco principale, sono stati presentati tutti i 13 stimoli uditivi (in ordine casuale) e viene chiesto ai soggetti di rispondere se hanno ascoltato / ba / o / va /, così come nel blocco di familiarizzazione (tre volte ciascuno in modo random). Per ogni soggetto quindi, viene selezionato il valore medio [μ] (lo stimolo che il

soggetto percepisce il 50% delle volte come / ba / o / va /) e la misura della soglia [σ] dove i soggetti percepiscono l'86% dei tempi chiaramente / ba / o / va /. Il valore di σ indica la pendenza della nostra funzione psicofisica (quando il valore è più piccolo, la curva è meno pianeggiante, ovvero la discriminazione delle due sillabe è più facile).

La misura della soglia tattile comprende due blocchi di familiarizzazione e un blocco principale. Viene chiesto ai soggetti soggetto di premere il tasto '1' ogni volta che si sente che percepiscono il movimento dei bottoni verso sinistra e premono il pulsante '2' quando è verso destra. Nel primo blocco di familiarizzazione, vengono presentati i bottoni verso sinistra con velocità diverse (-40ms, -30ms, -20ms, -15ms, -10ms, -5ms), ognuno ripetuto due volte (in ordine casuale), mentre nel secondo blocco di familiarizzazione vengono presentati i bottoni verso destra con velocità differenti (5ms, 10ms, 15ms, 20ms, 30ms, 40ms) ognuno ripetuto due volte ordine (in modo casuale). Per ogni soggetto quindi, viene selezionato il valore medio [μ] (lo stimolo che il soggetto percepisce il 50% del tempo come verso "sinistra" o verso "destra") e la misura della soglia [σ] in cui i soggetti percepiscono l'86% delle volte il movimento verso "sinistra" o verso "destra".

Dopo la misura delle soglie uditive e tattili (separatamente), sono state svolte le sessioni audio-tattili di familiarizzazione, 9 blocchi di training e 3 blocchi di test. In tutte queste fasi sono stati registrati nuovamente i valori medi e pendenza (per ciascun soggetto).

Nel blocco di familiarizzazione audio-tattile sono stati usati i tre stimoli uditivi che abbiamo usato per la misura della soglia uditiva (stimolo numero 1, 11, 21), in compresenza degli stimoli tattili che rispetto alla soglia tattile registrata assumessero valore di -3σ (associata allo stimolo uditivo 1), di + 3σ (associata allo stimolo uditivo 11) e di -3σ associata allo stimolo uditivo 21), presentati tre volte ciascuno in modo random. Successivamente ogni soggetto ha svolto 9 blocchi di training in cui veniva presentata sempre la stessa associazione audio-tattile, ovvero la sillaba /ba/ con il movimento in direzione sinistra e la sillaba /va/ con il movimento verso destra. Per tutti i blocchi sono stati utilizzati tutti e 13 gli stimoli uditivi, mentre le due stimolazioni tattili ad essi associati, cambiavano nei vari blocchi, da una stimolazione tattile sopra soglia ad una stimolazione tattile sotto soglia (meno chiara). In particolare si sono utilizzati: per il primo e secondo blocco +3σ dalla soglia tattile, terzo blocco + 2σ dalla soglia tattile, quarto blocco: +1.5σ dalla soglia tattile, quinto, sesto e settimo blocco + 1σ dalla soglia tattile, ottavo e nono blocco + 0.5σ dalla soglia tattile (si veda fig. 17).

Dopo le sessioni di training vengono svolti i tre blocchi di test, in cui viene chiesto ai soggetti di ignorare la stimolazione tattile (si spera in questo modo che i soggetti rispondano in base a ciò che ascoltano, in modo tale che il risultato ottenuto possa riflettere l'influenza tattile

non volontaria). Viene quindi detto ai soggetti di ignorare l'informazione tattile perchè questa non è attendibile e potrebbe essere fuorviante per dare le risposte (che vanno date invece, in base alla sillaba ascoltata di volta in volta).

Il primo blocco di test è identico all'ultima sessione della formazione audio-tattile (stimoli tattili: 0.5 σ dalla soglia tattile).

Nel secondo blocco di test, chiamato 'biased-weak' viene aggiunto un bias, ovvero + 3 al punto in cui i soggetti sentono il 50% del tempo 'ba' e 50 % dei tempi 'va' della soglia uditiva, ovvero abbiamo associato i 3 stimoli uditivi che venivano percepiti come /ba/ al posto di 3 stimoli che precedentemente erano percepiti come /va/ (si tenga presente che vengono sempre utilizzati tutti e 13 gli stimoli uditivi) e abbiamo associato agli stimoli uditivi due stimoli tattili sotto soglia (0.5 σ dalla soglia tattile) (si veda la fig. 18).

Nel terzo blocco di test, chiamato 'biased strong' viene aggiunto lo stesso bias del secondo blocco di test, ma stavolta le due stimolazioni tattili associate agli stimoli uditivi sono sopra soglia (5 σ dalla soglia tattile).

Infine, viene nuovamente misurata la soglia uditiva nello stesso modo in cui è stato fatto inizialmente. Ciò è stato fatto per verificare che la soglia uditiva fosse esattamente la stessa della misura iniziale, per escludere che i soggetti avessero avuto un eventuale over-learning delle sillabe uditive e per discriminare se i risultati ottenuti fossero frutto di eventuale effetto di interazione audio-tattile o piuttosto legati al un miglioramento della discriminazione uditiva.

Risultati.

E' stata svolta un'analisi t-test per dati appaiati tra i valori medi e la pendenza (o σ) della misurazione della soglia uditiva iniziale con i valori dei tre blocchi di test e della misurazione uditiva post esperimento. Come già detto, solo 8 soggetti (numero dei soggetti: 1, 2, 3, 7, 8, 9, 10, 14) sono stati inclusi nell'analisi (si veda fig. 22 in appendice).

Per quanto riguarda i valori medi, il t-test appaiato non ha mostrato differenze significative tra la soglia uditiva iniziale e il primo ($p = .166$), secondo o 'biased-weak' ($p = .247$) blocco di test e la misura della soglia uditiva post-esperimento ($p = .409$). É presente invece un effetto significativo per il valore medio della soglia uditiva iniziale e il terzo o 'biased strong' blocco di test ($p = .008$).

Per quanto riguarda i valori della pendenza (o σ), l'analisi t-test appaiato non ha mostrato alcuna differenza significativa. In particolare, non vi è alcuna differenza significativa tra il valore del sigma della soglia uditiva iniziale con il primo ($p = .762$), secondo ($p = .487$), terzo ($p = .269$) blocco di test e la soglia post-esperimento ($p = .909$). (si veda fig. 25)

La maggior parte dei soggetti (%) ha riportato un cambiamento della percezione delle sillabe dopo qualche blocco della sessione di training (si veda fig. 24 e discussione).

Discussione.

Il presente studio ha avuto come fine quello di verificare un'integrazione audio-tattile con stimoli linguistici semplici. Si è voluto verificare, quindi, l'informazione tattile possa modificare la percezione uditiva (in questo caso, la discriminazione di quale delle due sillabe venisse presentata). In particolare, ci si è chiesti se il tatto potesse influenzare la percezione delle sillabe anche quando venisse chiesto ai soggetti di ignorarne la presenza. Come condizioni di controllo, la presentazione degli stimoli tattili è stata svolta sia con stimoli sopra soglia che sotto soglia. Inoltre, si è verificato se la stimolazione tattile potesse influenzare la percezione uditiva anche quando fosse in contrasto con l'associazione precedentemente appresa (durante le sessioni di apprendimento).

I risultati mostrano che gli input tattili hanno influenzato la percezione uditiva solo quando erano chiari (test 'biased strong'), ovvero la stimolazione tattile ha influenzato la discriminazione uditiva delle nostre sillabe /va / e / ba / solo se era sopra soglia. In altre parole, la curva psicometrica non cambia in presenza della stimolazione tattile sotto soglia. (si veda un esempio della curva psicometrica fig. 23in appendice).

Questo risultato non conferma la nostra ipotesi. Infatti, l'assenza di risultati significativi quando la stimolazione tattile è sotto soglia potrebbe invalidare perfino il risultato significativo dell'influenza tattile sull'udito quando l'informazione tattile è sopra soglia. Infatti, si potrebbe obiettare che, anche se abbiamo trovato un effetto del nostro bias, ciò potrebbe essersi verificato perché i soggetti non hanno seguito la nostra istruzione circa l'ignorare l'informazione tattile (si ricordi che abbiamo detto loro che la stimolazione tattile poteva essere fuorviante). Detto in altro modo, si deve considerare la possibilità che ciò che si è trovato è legato al fatto che i soggetti erano ancora focalizzati sulla stimolazione tattile, in modo "cognitivo", e non hanno risposto in base a un cambiamento della percezione. Secondo la letteratura, quindi, si potrebbe dire che ciò che abbiamo trovato è stato il risultato di una cosiddetta "strategia di selezione" (Massaro, 1987) legato all'affidabilità della modalità tattile, anziché un puro effetto di integrazione.

Il fatto che non si sia riscontrata alcuna differenza significativa per i valori di pendenza potrebbe essere dovuta al fatto che i valori di sigma uditiva erano già troppo piccoli per trovare un valore migliore (ovvero la discriminazione delle sillabe è stata abbastanza facile in tutti i blocchi).

Per quanto riguarda gli effetti percettivi soggettivi, come accennato, la maggior parte dei nostri soggetti ha sperimentato una sorta di "percezione allucinatoria", ovvero dopo qualche blocco hanno cominciato a sentire suoni strani che non erano nemmeno riconducibili a sillabe. Fatta eccezione per pochi casi, questo effetto uditivo è scomparso nei post-test uditivi, non possiamo escludere il fatto che questo effetto non sia dovuto all'interazione audio-tattile o è legato esclusivamente alla fatica uditiva o all'adattamento.

Per verificare se si trattasse di un'influenza del dispositivo tattile o un effetto allucinatorio uditivo, si è svolta una prova con un soggetto naïve, utilizzando solo le cuffie. Il soggetto ha riportato l'effetto allucinatorio come per la maggior parte dei soggetti. Ragion per cui, ciò ha confermato che l'effetto percettivo uditivo non era dovuto all'integrazione audio-tattile, ma ad un cambiamento della percezione uditiva. In effetti, in letteratura vengono riportati i fenomeni di adattamento uditivo in presenza di più sillabe ripetute, denominato 'adattamento selettivo'(riportato con le sillabe /aba/ e /ada/ da Bertelson e colleghi, nel 2003)

Conclusione

Si è svolto questo esperimento di soglie psicofisiche, con l'intento di verificare e riscontrare prove a favore dell'integrazione audio-tattile, nonché dell'influenza della stimolazione tattile sulla percezione uditiva di stimoli linguistici, anche quando la stimolazione tattile è sotto soglia.

Come si è visto, dal momento che i nostri soggetti non hanno riportato un cambiamento della percezione uditiva in presenza di stimolazione tattile sotto soglia ma solo quando la stimolazione tattile era sopra soglia, non si può escludere che l'effetto ritrovato non sia dovuto a strategia di selezione, piuttosto che ad un vero e proprio effetto percettivo di integrazione audio-tattile. Una ipotesi plausibile è che il tatto per influenzare la percezione uditiva debba essere necessariamente sopra soglia. Questo argomento sarebbe a favore della nostra ipotesi di una integrazione audio-tattile anche per stimoli linguistici, ma come detto, poiché la stimolazione tattile è affidabile (al di sopra della soglia), le risposte potrebbero essere date seguendo tale informazione, ignorando le istruzioni.

La presenza dell'effetto allucinatorio percettivo aveva inizialmente incrementato le nostre speranze. Infatti, in letteratura sono presenti casi in cui l'effetto di integrazione di due modalità comporta la 'creazione' percettiva di stimoli differenti da quelli presentati originariamente nelle due modalità. Sebbene presente quindi in letteratura come fenomeno, la mancata conferma che tale fenomeno potesse essere dovuto in questo esperimento per via di una integrazione audio-tattile (dal momento che si verificava anche in assenza di stimolazione

tattile) non ha corroborato la nostra ipotesi iniziale, non portando ad una prova in favore di una integrazione audio-tattile a livello percettivo.

2.2.1 Pilota alla ricerca dell "effetto McGurk" audio-tattile

Come abbiamo visto nella discussione e conclusione dell'esperimento psicofisico di discriminazione delle sillabe /ba/ e /va/, non è stata confermata l'integrazione audio-tattile e l'influenza del dispositivo tattile sulla percezione uditiva come un puro "fenomeno percettivo". Guardando alla letteratura in merito a fenomeni percettivi puri, si è considerato l'effetto McGurk come un buon esempio di una integrazione crossmodale pura, ovvero come un fenomeno di tipo automatico o puramente percettivo. In breve, l'effetto McGurk originale (McGurk & MacDonald, 1976) mostra un'integrazione percettiva tra la modalità uditiva e quella visiva. Nell'effetto originale la presentazione visiva ripetuta della sillaba /ga/ doppiata in simultanea con la presentazione uditiva ripetuta della sillaba /ba/ comporta la percezione finale della sillaba /da/. Inoltre, la presentazione visiva della sillaba /ba/ in simultanea alla presentazione uditiva della sillaba /ga/ viene talvolta percepita come /bga/.

In letteratura, sono presenti alcuni tentativi di riproporre un equivalente dell'effetto originale McGurk (Fowler e Dekle, 1991; Sato et al. 2010) nel dominio audio-tattile, ma i risultati sono stati contrastanti e non sistematici, come per il McGurk originale.

Si è cercato quindi, di creare una condizione similare al McGurk originale, come prova in favore di una vera integrazione audio-tattile. In altre parole, si è cercato di rispondere alla nostra domanda circa la possibilità che la stimolazione tattile possa influenzare effettivamente la percezione uditiva, traendo 'ispirazione dall'effetto del McGurk originale, traducendolo in un effetto McGurk audio-tattile.

Si è svolto uno studio pilota di tipo qualitativo. Dal momento che la stimolazione uditiva fornita dal nostro dispositivo è totalmente arbitraria, si è chiesto a due soggetti di svolgere un periodo molto lungo di addestramento, in modo da poter implicare una forte associazione tra ogni sillaba con uno specifico pattern tattile, in modo similare allo studio psicofisico sopra esposto, ma stavolta utilizzando quattro sillabe associate a quattro pattern di movimento diversi. La scelta di usare quattro sillabe (registrate più volte) è stata fatta per evitare l'effetto allucinatorio riscontrato nel precedente esperimento. Anche in questo caso sono state scelte le sillabe in base alla matrice di confusione di consonanti, di Miller and Nicely (1955) (si veda fig. 3 in appendice) scegliendo quattro sillabe facilmente confusi, per altro includenti quelle usate negli studi del McGurk originale.

Inoltre, sono state svolte delle piccole sessioni di test (i soggetti controllano le loro risposte dopo averle date), in cui venivano presentati anche degli stimoli incongruenti rispetto al training (si veda procedura). In modo similare al McGurk originale, ci si aspetta di trovare in presenza di stimoli incongruenti rispetto al training, un effetto percettivo di integrazione e di 'creazione' di una sillaba differente rispetto a quella udita. .

Metodo

Soggetti

Hanno svolto lo studio pilota 2 soggetti maschi (età media 35), uno di madrelingua russa e l'altro di madrelingua italiana.

Strumenti

Sono stati utilizzati un comune pc portatile, le stesse cuffie dello studio precedente 'Marshall Monitor' e il dispositivo chiamato *FeelSpeech*. Il dispositivo *Felspeech* presenta quattro diverse stimolazioni tattili, ovvero il movimento (chiamato *microswiping*) di quattro file di bottoni della matrice 4x4 in quattro direzioni (in alto, in basso, a destra e a sinistra). (si veda fig. 4 b) e c) in appendice).

Stimoli

Stimoli uditivi: sono state selezionate 4 sillabe /ba/, /va/, /ga/, /da/, registrate 10 volte ciascuna (40 stimoli complessivi) per evitare l'effetto uditivo di adattamento, già citato. Gli stimoli uditivi sono sempre chiari (o sopra soglia), dal momento che per nel McGurk originale, tutti gli stimoli devono essere chiari (o sopra soglia).

Stimoli tattili: sono stati utilizzate quattro direzioni di movimento del *microwiping*, ciascuna delle quali viene associata ad una sillaba. Dal momento che per nel McGurk originale, tutti gli stimoli devono essere chiari (e che nello studio precedente, il tatto ha avuto un effetto significativo solo quando presentato sopra soglia) tutti gli stimoli sono stati presentati in modo chiaro (o sopra soglia). (si veda fig. 6 in appendice)

Stimoli audio-tattili: tutte le sillabe audio /ba/ sono state associate al movimento della prima fila di bottoni verso il basso, tutte le sillabe audio /va/ sono state associate al movimento dell'ultima fila di bottoni verso l'alto, tutte le sillabe audio /ga/ sono state associate al movimento della fila di bottoni in basso verso destra, tutte le sillabe audio /da/ sono state associate al movimento della fila di bottoni in alto verso sinistra (si veda fig. 5).

Procedura

Ciascun soggetto ha svolto giornalmente varie ore di apprendimento (pausa dopo ogni sessioni di allenamento) e un test a fine giornata. Complessivamente, il soggetto 1 ha svolto 6 giorni di allenamento per 4-5 ore al giorno (circa 50.000 prove), mentre il soggetto 2 ha svolto 6 giorni di allenamento per 1,5-2 ore al giorno (circa 15.000 prove).

Sia nei blocchi di allenamento che nel test, viene chiesto ai soggetti di premere dei pulsanti della tastiera QWERTY in base alla sillaba percepita (facendo riferimento alla sillaba udita, sebbene nel training la stimolazione audio-tattile è sempre congruente per rafforzare le associazioni). Si tenga presente che, mentre il Soggetto1 rispondeva in base al alla localizzazione dei bottoni (non considerando, quindi, la direzione del movimento dei bottoni): W (per DA), A (per BA), S (per VA), D (per VA); il Soggetto2 rispondeva in base al movimento dei bottoni: W (per VA), A (per DA), S (per BA), D (per GA). (si veda fig. 5 e 6) (si veda fig 7 in appendice).

Le sessioni di allenamento, composte da 200 prove ciascuna, sono state sia "attive" che "passive".

Nelle sessioni di allenamento "passive", il computer presenta un nuovo stimolo (una sillaba delle quattro, con la sua stimolazione tattile associata, in modo casuale) ogni volta che il soggetto risponde in base alla sillaba che ha ascoltato (tranne per la condizione ‘nessuna risposta, spiegata in seguito). Queste sessioni includono tutti gli stimoli uditivi (40 sillabe) presentati in modo random in quattro condizioni: solo uditiva, solo tattile, audio-tattile e nessuna risposta (dove viene chiesto ai soggetti di pensare quale sillaba sia senza premere nessun tasto). Inoltre, in una prova su cinque (un quinto delle intere sessioni) veniva presentata solo la stimolazione uditiva per invitare i soggetti a focalizzarsi sull’informazione uditiva e non su quella tattile (in altre parole a costringerli a non fare affidamento solo sulla stimolazione tattile e prestare attenzione anche alla stimolazione uditiva).

Nelle sessioni di allenamento “attive”, i soggetti leggevano la sillaba mostrata sullo schermo mentre veniva presentata la stimolazione tattile. Anche le prove attive presentano le quattro condizioni citate sopra.

Il test, effettuato una volta al giorno a fine addestramento, è composto da 240 prove. Il test presenta uno squilibrio delle prove congruenti e incongruenti rispetto al training (in una proporzione di 10:3 prove congruenti vs. incongruenti) al fine di evitare invalidare le associazioni apprese dei partecipanti. Il test quindi, comprende sequenze di 10 prove congruenti (es. / BA / presentate 10 volte con la stimolazione tattile appresa) e tre sequenze di prove incongruenti (ogni sillaba associata a ciascuna delle possibili incongruenti tattile

stimolazioni) ordine. Per ottenere gli stimoli, i soggetti premono i tasti sulla tastiera, ma per verificare l'esistenza o meno di una 'illusione percettiva' (come atteso nelle prove incongruenti, di cui il soggetto è ignaro dell'ordine di apparizione), ogni stimolo viene presentato 10 volte (sempre lo stesso), così che i soggetti prima della fine delle 10 presentazioni, alzano il pollice dal dispositivo per controllare la stimolazione tattile effettiva e verificare se la percezione uditiva corrispondesse con la stimolazione tattile o no. Come prove incongruenti sono state presentate tutte le 12 combinazioni incongruenti possibili di della combinazione degli stimoli uditivi e tattili.

Risultati e Discussione

Trattandosi di uno studio qualitativo, i dati ottenuti sono solo quelli riportati dai soggetti. Ci si attendeva che nelle prove incongruenti i soggetti riportassero una percezione uditiva differente, o influenzata, dalla stimolazione tattile, ed invece i soggetti non hanno riportato nessun cambiamento della percezione uditiva e quindi nessun equivalente del McGurk originale. Entrambi i soggetti hanno riferito che dopo l'esperimento hanno smesso di sperimentare l'effetto McGurk, precedentemente esperito normalmente. Se da un lato, quindi, una critica che può essere fatta a questo esperimento è quella di aver utilizzato stimoli uditivi e tattili non ottimali per esperire un simil McGurk audio-tattile (che si ricordi, ancora non presente in letteratura come effettivamente un fenomeno automatico per la maggior parte della popolazione), d'altro canto, il fatto che i soggetti hanno smesso di esperire il McGurk classico può portare a ipotizzare un effetto dell'allenamento, per così dire, 'eccessivo' a prestare attenzione alla modalità uditiva nella percezione delle sillabe allenate (tra cui, appunto quelle del McGurk originale). Inoltre, in assenza di prove su un effetto McGurk audio-tattile, non possiamo escludere l'ipotesi che un effetto McGurk audio-tattile possa funzionare in un modo diverso da quello audiovisivo, che coinvolge altre sillabe, sia utilizzando altri tipo di stimolazione tattile, o un periodo di addestramento maggiore.

Ad ogni modo, ciò non toglie, che ancora una volta, non si è riusciti a trovare una prova per una integrazione audio-tattile a livello percettivo, utilizzando stimoli linguistici.

Conclusione in merito all'integrazione audio-tattile con stimoli linguistici.

Nell'introduzione di questa sezione si è sottolineato che diversi studi sull'integrazione multisensoriale hanno dimostrato che l'integrazione audio-tattile è comune, ma che al momento

non è ancora presente in letteratura una prova chiara circa il fatto che l'integrazione audio-tattile possa verificarsi all'interno di un discorso, o con stimoli linguistici.

Sia lo studio psicofisico della discriminazione di sillabe che lo studio pilota alla ricerca dell'effetto McGurk audio-tattile non hanno corroborato l'integrazione audio-tattile come fenomeno puramente percettivo con stimoli linguistici

Nel primo studio, come detto, ci si aspettava di trovare un'influenza della modalità tattile sulla percezione e discriminazione di sillabe, con stimoli tattili sia sotto che sopra soglia. Come si è detto, il fatto che si sia riscontrata una influenza tattile nella discriminazione di sillabe solo quando era sopra soglia, non comporta un' inequivocabile influenza della modalità tattile su quella uditiva, dal momento che i soggetti possono aver risposto in base alla modalità 'disambiguante', riportando le risposte 'sentite tattilmente' e non 'ascoltate'. Per la verità, questa è una interpretazione critica che può essere fatta in diversi studi e che è riconducibile alla questione comune dell'interpretazione degli studi crossmodali circa un l'effettivo cambiamento della strategia/criterio nel dare le risposte (Massaro, 1987).

Inoltre, è stato anche ipotizzato che il mancato risultato atteso possa essere stato legato ai tempi di apprendimento, ma anche il pilota che presenta un periodo di apprendimento prolungato, non ha dato conferma a questa ipotesi. L'interpretazione più estrema che può essere fatta in merito al fatto che forse questa integrazione non ha avuto luogo perchè dovuta a tempi di apprendimento non sufficientemente lunghi per stabilire un'associazione che porti ad una 'automaticità' dell'associazione audio-tattile, è che forse i fenomeni di integrazione crossmodale richiedono che questa associazione sia intrinseca (avvenga in modo sistematico), o sia appresa sin dalla nascita.

Rimane così da chiarire, sia il ruolo del criterio scelto, sia dei tempi di apprendimento, sia del tipo di informazione utilizzata. Infatti, non è da escludere anche una terza ipotesi, riguardante il fatto che si sono utilizzati come stimoli acustici, delle sillabe, ovvero stimoli che presentano una complessità (o se vogliamo, contenuto di frequenze maggiore, per voler citare la letteratura REF) rispetto alla complessità dell'informazione tattile.

In questa sede, non viene fatto riferimento al ruolo della semantica, dal momento che si suppone che le sillabe presentate in modo continuo, perdono in parte il valore semantico e possono essere 'trattate come' dei suoni (si vedano i risultati riportati e il fenomeno di abituazione), anche perchè isolate da un contesto che le possa disambiguare.

2.2.4 Studio pilota con due toni.

Per completezza, riporto un piccolo pilota che è stato svolto sempre con l'intento di trovare una prova dell'integrazione audio-tattile. Dal momento che non è stato riscontrato alcun effetto con gli stimoli linguistici, si è pensato di svolgere un pilota, utilizzando gli stimoli uditivi tonali, per i quali è stato riscontrato un effetto di integrazione audio-tattile in letteratura (per esempio Schürmann e coll., 2004).

Riprendendo il compito dell'esperimento per la detezione delle incongruenze temporali in flussi ritmici, viene chiesto ad un soggetto di rilevare, in una presentazione ritmica,quando due suoni e due stimoli tattili non sono congruenti rispetto ad un'associazione audio-tattile prestabilita. Viene manipolata la velocità del flusso per vedere se ad alta velocità il soggetto riesce a individuare le incongruenze audio-tattili. Questo pilota quindi, da un lato mira ad analizzare cosa accade in presenza di stimolazione audio-tattile incongruente rispetto ad un'associazione imparata (e come si vedrà, riferita 'automatica e naturale' dal soggetto musicista) e dall'altro, a vedere entro quale finestra temporale minima, il soggetto riesce a rilevare le incongruenze. Come detto, quindi, gli obiettivi di tale pilota sono similari ai due precedenti studi, sia riferiti all'analisi delle finestre temporali che di integrazione audio-tattile, stavolta con stimoli non linguistici. Ci si aspetta che il soggetto nelle prove incongruenti possa esperire un suono diverso rispetto ai suoni presentati e che a finestre temporali piccole (100ms) non riesca a individuare le incongruenze.

Metodo

Soggetto

Ha preso parte allo studio pilota un soggetto (femmina, 29 anni, sperimentatore). Il soggetto studia e suona il violino da 20 anni.

Strumenti

Sono stati utilizzati un comune pc portatile, le stesse cuffie dello studio precedente 'Marshall Monitor' e il dispositivo chiamato *FeelSpeech*. Il dispositivo *Felspeech* è a due canali, ovvero due celle di matrice 2x4 ciascuna. Quando presente lo stimolo tattile, tutti i bottoni di ciascuna cella si alzano. (si veda fig. 4 d) in appendice).

Stimoli

<u>Stimoli uditivi.</u> sono stati presentati 2 toni da 300 Hz e 4000 Hz.

<u>Stimoli tattili.</u> gli stimoli tattili prevedevano la presenza di tutti i bottoni di ogni cella alzati.

Entrambi gli stimoli sono stati presentati alle velocità di 300 ms, 250 ms, 200 ms, 150 ms e 100 ms (la durata degli stimoli e delle pause tra uno stimolo e l'altro era la medesima).
Stimoli audio-tattili: il suono grave è stato associato al dito anulare e il suono acuto al dito indice della mano sinistra.
La durata degli stimoli e delle pause tra uno stimolo e l'altro sono identiche, dal momento che in un pre-pilota utilizzando pause di durata differente (di 100 ms, 200 ms, 300 ms, 400 ms con gli stimoli di durata di 150 ms) il compito era risultato troppo complesso da svolgere. Inoltre, nel pre-pilota è stata provata anche l'associazione 'inversa' degli stimoli audio-tattili (suono acuto con l'indice e suono grave con l'anulare della mano sinistra), ma il soggetto ha riferito che fosse un'associazione del tutto 'innaturale' e 'quasi fastidiosa', per cui si è deciso di svolgere l'associazione più 'naturale' (dal momento che il nostro obiettivo era di valutare un effetto percettivo di integrazione nelle prove incongruenti, e il soggetto doveva apprendere un'associazione, si è pensato di utilizzare un'associazione che fosse più facilmente apprendibile, ancora meglio se già familiare al soggetto).

Procedura

Il soggetto ha svolto 15 sessioni in quattro giorni. Ogni sessione prevede 1 minuto di familiarizzazione, 3 esempi di incongruenze e 10 minuti di test. Viene chiesto al soggetto di premere la barra spaziatrice ogniqualvolta rileva delle incongruenze tra l'associazione audio-tattile prestabilita. Come accennato, è stata manipolata la velocità degli stimoli e del loro susseguirsi (si veda la sezione 'stimoli').
Il primo giorno il soggetto ha svolto 5 sessioni di 150 ms per verificare la velocità di miglioramento della performance alla stessa velocità.
Il secondo giorno si sono manipolate le velocità per valutare la performance con diverse finestre temporali e il soggetto ha svolto 3 blocchi da 300 ms, 250 ms e 200 ms.
Il terzo giorno il soggetto ha svolto altri 5 blocchi: da 250 ms, 200 ms, 150 ms, 100 ms e 150 ms.
Dal momento che a velocità 100 ms la performance era molto bassa (si veda risultati), il quarto giorno il soggetto ha svolto altri 2 blocchi da 150 ms, per valutare fino a che punto la performance potesse migliorare.
In un secondo momento, dopo l'esperimento, si è svolto un 'test' per valutare la grandezza dell'effetto della percezione 'del suono intermedio' riportata dal soggetto (si veda risultati e discussione). Nel test, viene chiesto al soggetto di premere la barra spaziatrice ogni volta il 'tono intermedio'. Il soggetto non viene reso a conoscenza del fatto che tutte le prove

incongruenti fossero state presentate con un unico tipo di incongruenza audio-tattile (si veda discussione).

Risultati e discussione

Come accennato nell'introduzione dello studio pilota, si sono cercati di valutare e verificare due ipotesi. Da un lato, si è voluto riverificare entro quale margine temporale si potesse individuare un'incongruenza audio-tattile, e quindi, un appaiamento tra stimoli appartenenti alle due modalità (finestra temporale crossmodale). Dall'altro, si è cercata nuovamente la prova per un effetto puramente percettivo di integrazione audio-tattile, sperando che nelle prove incongruenti si potesse verificare qualcosa a livello percettivo (come per esempio, udire un suono diverso dai due suoni presentati).

Riguardo al primo punto, i risultati mostrano che la performance migliore ottenuta dal soggetto si è verificata quando entrambi gli stimoli (e le loro pause) erano a 150 ms (si veda fig.). Questo risultato è stato stabilito guardando alla performance del terzo giorno. Dal quarto giorno in poi, si è deciso di svolgere altri blocchi con stimoli a 150 ms, per valutare la nostra seconda ipotesi, ovvero, cercare di valutare un'eventuale cambiamento percettivo quando il soggetto era 'cosciente' delle incongruenze. Gli ulteriori blocchi a velocità 150 ms ci hanno permesso anche di valutare quando la performance ad una stessa velocità può migliorare (come per il primo giorno di training), e si è visto che rimane pressoché costante, il che potrebbe rispondere all'obiezione che qualcuno potrebbe fare in merito al fatto che il soggetto possa aver ottenuto una performance migliore a 150ms dal momento che il primo giorno è stato 'allenato' al compito a questa velocità. Inoltre, se si guarda anche alla performance delle finestre temporali svolte in precedenza, stimoli crossmodali a 100 ms sembrano essere particolarmente difficili da rilevare. Ovviamente ciò potrebbe essere ulteriormente confermato risvolgendo l'esperimento con un soggetto naive, senza la sessione di training a velocità costante, oppure utilizzando velocità diversa nel training iniziale (per es. a 100 ms) e vedere se anche dopo la manipolazione delle velocità si ottiene una performance migliore alla velocità 'più allenata'.

Per quanto riguarda la nostra seconda ipotesi, circa l'eventuale presenza di una percezione differente nelle prove incongruenti, si è verificata a livello percettivo, ma non nel test che si è svolto dopo l'esperimento. Come detto, si è scelto di usare l'associazione audio-tattile definita come 'naturale dal soggetto, per velocizzare i tempi di apprendimento ed il verificarsi eventuale dell'effetto percettivo. Tale associazione, ha riportato il soggetto, rifletteva la disposizione dei suoni nella tastiera del violino per cui i suoni più acuti vengono prodotti utilizzando il terzo dito (anulare) rispetto al primo dito (indice) della mano sinistra. Il soggetto,

dopo qualche blocco di training, ha riportato di 'ascoltare' un suono intermedio ogni volta che veniva stimolato il dito indice (normalmente associato con il suono grave) in compresenza del suono acuto. Per l'esattezza, ha specificato che trovava il suono acuto sgradevole e che il suono intermedio esperito era 'più vicino' al suono acuto rispetto al suono grave. Non si verificava questa sensazione in presenza dell'incongruenza opposta (suono grave presentato insieme alla stimolazione del dito anulare).

Dal momento che il soggetto ha riportato questa sensazione, che sembrava per altro, essere sistematica, si è deciso di svolgere un test, in cui veniva chiesto di premere la barra spaziatrice ogni volta che esperiva il suono intermedio, ma in realtà tutti gli stimoli erano incongruenti. Nel test, il soggetto non è riuscito a capire che tutte le prove incongruenti presentavano lo stesso tipo di incongruenza e ha mostrato l'effetto percettivo a meno del 30% delle volte in tutte e tre le sessioni di test svolte (si veda fig). Questo risultato dei test non è quindi riuscito a confermare a livello quantitativo, l'effetto riportato percettivamente dal soggetto. In altre parole, si è verificato che questo effetto percettivo non avviene in modo sistematico. Ragion per cui, non è stato possibile confermare l'entità del fenomeno con la tipologia di test utilizzato. Una delle considerazioni che può essere fatta in merito, è che per far si che l'effetto percettivo si verifichi, occorre che il soggetto sia consapevole della stimolazione incongruente. Dal momento che il soggetto non è riuscito a capire che l'incongruenza presentata era sempre la stessa, si potrebbe ipotizzare che il 30% dell'effetto rilevato dal test, rifletta solo le prove in cui il soggetto ha rilevato l'incongruenza audio-tattile che comportava l'effetto percettivo. Più oscuro rimane il perché il soggetto abbia riferito un effetto percettivo solo in presenza di una particolare tipologia di incongruenza della stimolazione audio-tattile e non dell'altro tipo di incongruenza. Un'ipotesi che si può fare in merito potrebbe riguardare il ruolo della salienza degli stimoli. Il soggetto infatti, riporta come 'fastidioso' il suono acuto, e 'difficile da ignorare', rispetto al suono grave.

Interpretazioni in merito richiederebbero ulteriori indagini con un test sufficientemente sensibile per la misurazione quantitativa dell'effetto percettivo riscontrato qualitativamente.

2.3 Sostituzione sensoriale: 'ascoltare' con il tatto

Come abbiamo visto nella parte introduttiva, una parte della ricerca che ipotizza che le modalità sensoriali sono 'funzionalmente' accomunabili, ovvero, che potenzialmente ogni modalità sensoriale può veicolare (e quindi 'contiene' un corrispettivo) dell'informazione veicolata da un'altra modalità, attraverso un altro codice. Ci si è chiesti quindi, se sia possibile veicolare informazione uditiva tramite modalità tattile. Nello specifico ci si è focalizzati sugli studi che hanno utilizzato la modalità tattile per veicolare informazione acustica di tipo linguistico. Come visto, la letteratura riporta dati contrastanti e non completamente convincenti circa i sistemi di sostituzione sensoriale tattili che sono stati utilizzati fino ad ora per aiutare persone sorde o con deficit uditivi a comprendere il linguaggio. Per lo più, questi strumenti sono stati abbandonati perché richiedevano un grande sforzo nell'apprendimento, non erano portatili e dopo un pò creavano un effetto di 'formicolio'. Ci si è chiesti quindi, se una volta superati questi limiti si possa verificare, e ritrovare, la facilitazione della modalità tattile sulla comprensione linguistica, in presenza di stimoli uditivi di tipo linguistico, non completamente chiari. Si è quindi, cercato di superare i limiti dei precedenti sistemi di sostituzione sensoriale con il *FeelSpeech*. Nello specifico, il limite della portabilità è stato superato costruendo una versione del dispositivo strumento piccola e maneggevole, il limite del formicolio è stato superato grazie al '*microswiping*' (si veda descrizione dello strumento) che risulta essere una stimolazione piacevole e non comporta il fenomeno della 'desensibilizzazione' (non essendo né elettrica né vibratoria).

Come citato, chiarire se la modalità tattile può influire la percezione uditiva, oltre ad essere un elemento chiave nella comprensione della percezione crossmodale, circa il ruolo delle modalità sensoriali, potrebbe avere fini applicativi utili con le popolazioni con deficit uditivi. Il progetto a cui ho preso parte presso l'Università Descartes di Parigi, seguita dalla

supervisione del prof. O'Regan, ha come obiettivo quello di poter aiutare questa popolazione a 'sentire meglio', utilizzando l'informazione tattile, in aggiunta alle informazioni uditive mancanti, così come accade nella popolazione di soggetti con principio di sordità o deficit uditivi.

Per tale fine sono stati svolti degli studi pilota e uno studio vero e proprio.

Sono stati svolti due studi pilota in cui si è cercato di 'ascoltare il tatto senza informazione uditiva'. In particolare, il primo studio pilota ha avuto come fine quello di verificare la capacità di discriminazione tattile di per sé, ad una velocità simile a quella del parlato (in assenza di informazione uditiva). Il secondo studio pilota, anche esso, senza informazione uditiva, è stato volto sempre a verificare l'utilità del dispositivo tattile per seguire meglio l'informazione labiale, in una situazione simile al cocktail party, cercare di verificare l'impressione di ascoltare il parlante, tramite l'informazione tattile e la lettura delle labbra.

Inoltre, si è svolto uno studio e uno studio pilota per 'ascoltare il tatto con parziale (o residua) informazione uditiva', ovvero, si è tentato di confermare l'ipotesi della sostituzione sensoriale di esperire di 'ascoltare' meglio un audiolibro con informazione uditiva degradata, tramite il tatto.

2.3.1 'Ascoltare' con il tatto, senza informazione uditiva.

2.3.1.1 Pilota con il linguaggio tattile.

Finora si è parlato della possibilità e della ricerca di evidenze in merito alla possibilità che un sistema di sostituzione sensoriale possa facilitare la percezione di informazioni parzialmente mancanti, di un'altra modalità sensoriale. Prima di procedere a verificare un fenomeno di sostituzione sensoriale vero e proprio, poiché uno dei problemi dei dispositivi tattili è la veicolazione delle informazioni tattili alla velocità normale di conversazione, si è pensato di svolgere un pilota per valutare la capacità di 'decodifica' tattile alla velocità normale del linguaggio. Per far ciò, si è pensato di allenare un soggetto in un compito 'simil-ecologico', utilizzando un linguaggio tattile continuo con la velocità normale del parlato. Dal momento che spesso, durante l'esposizione ad una nuova lingua, in uno stato iniziale si comprendono solo poche parole e man mano si incrementa il vocabolario del parlante, si è deciso di allenare il soggetto a riconoscere delle 'simil-parole' tattili, ovvero di riconoscere determinate sequenze tattili, all'interno di un discorso tattile (manipolando la velocità del testo e la quantità di parole/sequenze) con un lungo periodo di training.

Ci si aspetta che dopo un lungo periodo di training, il soggetto possa riconoscere le parole all'interno di un linguaggio alla velocità naturale del parlato, senza implicare uno sforzo

cognitivo. Se infatti, il riconoscimento delle parole può avvenire senza uno sforzo cognitivo (con poco training e solo con un'esposizione al linguaggio quotidiana), si risolverebbe uno dei motivi principali dell'abbandono dei dispositivi tattili: il grande sforzo cognitivo richiesto per il loro efficace utilizzo .

Metodo

Soggetti

Ha preso parte allo studio pilota, un soggetto di sesso femminile (età 15 anni). Il soggetto è bilingue, di madrelingua inglese e francese. Prima di ogni sessione sperimentale ha firmato il consenso informato. Il soggetto è stato pagato 10 euro l'ora (*ERC PoC Grant FeelSpeech*).

Strumenti

Sono stati utilizzati un comune pc portatile, le stesse cuffie dello studio precedente 'Marshall Monitor' (in questo caso per isolare acusticamente il soggetto) e il dispositivo chiamato *FeelSpeech*. Il dispositivo *Felspeech* presenta 12 diverse stimolazioni tattili associate a 15 sillabe diverse (l, b, p/dh, v, d, f, m/zh, g, s, z, k, t/ th, si veda fig.,8 a)), ovvero il movimento chiamato *microswiping*, in 12 direzioni differenti (combinazioni dei movimenti verso l'alto, basso, destra, a sinistra e in diagonale), della matrice di bottoni 4x4. (si veda fig 4 b) e c)).

Stimoli

Stimoli tattili: ciascuna delle 15 sillabe (alcune sono accomunate, in base alla fonetica del vocabolario inglese) viene presentata con una stimolazione tattile specifica. E' stato chiesto al soggetto di individuare 4 non-parole formate da 4 sillabe (4 movimenti) ciascuna ('pabalava', 'dagafama', 'pagadaba', 'vasalava') all'interno di un testo 'solo tattile'. (si veda fig. 8 b)).

Sono stati utilizzati due testi: il testo della favola di Biancaneve, che conteneva la parola "pabalava" localizzata casualmente nel testo, utilizzato per il primo training e il testo della fiaba di Alice nel paese delle meraviglie, che conteneva tutte e quattro le parole casualmente nel testo (si veda fig. 8 a) e b)). La favola di Biancaneve includeva 38 volte la parola 'pabalava'. La fiaba di Alice nel paese delle meraviglie includeva 21 volte la parola 'pabalava', 20 volte la parola 'dagafama', 21 volte la parola 'pagadaba' e 21 volte la parola 'vasalafa'. (si veda fig. 9 a) e b)).

La velocità di presentazione delle parole, come accennato è stata manipolata. Sono state utilizzate le velocità a 40, 35, 30, 25, 20, 15, 10, che indicano la velocità di presentazione dei fonemi, in modo convenzionale (40 indica velocità 4 volte più lenta del parlato). La frequenza

normale del parlato è indicata con ‘velocità 10’, che indica che vengono presentate 120 parole al minuto.

Procedura

Il soggetto ha svolto un periodo di training e test di un'ora o un’ora e mezza al giorno per un totale di 17 giorni (il primo giorno è stata svolta una fase di familiarizzazione con il compito). Per ogni sessione, il soggetto svolgeva più sessioni di training e talvolta veniva verificato con un test, la performance. Allenamento e il test, infatti, sono identici, eccetto che mentre nel training il soggetto può guardare lo schermo che visualizza ogni volta che compaiono le parole nel testo, durante il test, viene costretto a girarsi e a riferire la parola che ha sentito appena la sente (e lo sperimentatore verifica a computer l’effettiva comparsa della parola). Durante l'allenamento il soggetto dice "sì" ogni volta che crede di aver sentito le parole (serve a tenere il soggetto attivo durante la fase di allenamento). Come accennato, durante il test non ha la possibilità di guardare lo schermo e deve dire le parole ogni volta che le sente tattilmente, nominando le parole. In questo modo lo sperimentatore verifica se le varie parole sono state percepite o meno, o se sono state confuse tra di loro. Come accennato, il testo di "Biancaneve" è stato utilizzato solo per cinque giorni di allenamento su 17 sessioni complessive. E’ stato richiesto al soggetto di riconoscere tutte e 4 le non-parole, partendo da una sola sequenza tattile (una parola di 4 sillabe), presentata prima lentamente e poi sempre più velocemente. Ogni volta che veniva raggiunto il 60% di accuratezza nell’individuare una parola (misurata nel test), ne veniva aggiunta una nuova parola, fino all’apprendimento di tutte e 4 le non-parole (o sequenze tattili). Ovviamente, prima di aggiungere una parola nuova, veniva presentata singolarmente per familiare con la sequenza tattile corrispondente alla parola (si veda fig.)

Risultati

Come si vede dalla procedura (si veda fig), è stata aggiunta la seconda parola 'dagafama', il terzo giorno di allenamento (dopo aver raggiunto il 58% di accuratezza delle risposte corrette nel riconoscere 'pabalava' in 'Biancaneve' e il 71% in 'Alice nel paese delle meraviglie' alla velocità normale del parlato). E’ stata aggiunta la terza parola "pagadaba" il settimo giorno di allenamento (dopo aver raggiunto il 64% di accuratezza con 'pabalava' e 67% con 'dagafama' alla metà della velocità normale della conversazione, 20). E’ stata aggiunta la quarta parola "vasalafa" nella dodicesima sessione di allenamento (dopo aver raggiunto il 91% per 'pabalava', 62% per 'dagafama', 62% per 'pagadaba' a velocità 25).

In conclusione, dopo l'intero periodo di allenamento, nella fase di test del'ultima sessione il soggetto ha raggiunto la sua performance migliore alla velocità 15 (simile alla velocità del parlato normale): 82% per 'pabalava', 33% per 'dagafama', 67% per 'pagadaba' e 6% per 'vasalafa' . Mentre, a velocità 20 (metà velocità normale del discorso) ha ottenuto: 82% per 'pabalava', 48% per 'dagafama', 71% per 'pagadaba' e 29% per 'vasalafa'.(si veda tab.1 e fig. 26 in appendice).

Discussione e conclusione

L'obiettivo di questo pilota è stato quello di verificare l'utilizzabilità del dispositivo in presenza di una stimolazione tattile alla velocità normale del parlato.

Dopo 16 ore di addestramento, il nostro soggetto ha imparato a riconoscere quattro non-parole composte da 4 sillabe in un contesto simil-ecologico di un linguaggio 'tattile'. Come visto, il soggetto è stato in grado di riconoscere oltre il 60% solo tre parole su quattro a velocità vicina a 'tre volte meno' rispetto alla normale velocità del parlato.

Incoraggiante è stato ciò che ha riportato a livello percettivo, ovvero di aver sentito le parole come un unico 'chunk', ovvero come un'unica parola e non come una sequenza di input tattili e di riconoscere le parole senza sforzo cognitivo, ma solo "sentendoli" (ha avuto questa sensazione con tre parole su quattro, anche se anche la prestazione della seconda parola non è stata brillante). Ha riferito inoltre, che la quarta parola è stata più difficile da riconoscere (come è visibile guardando la scarsa performance con questa parola). Si può notare che, a differenza delle altre parole, quest'ultima presenta un pattern in diagonale, il che può far pensare che o è stato un pattern poco allenato o meno chiaro rispetto ai pattern in verticale e orizzontale, presenti in tutte le altre tre parole (si veda la fig.). Inoltre, ha riferito di trovare la consonante 'p' facile da riconoscere, il che può fare ipotizzare che vi possano essere dei pattern tattili più salienti di altri. Si tenga presente che parte della salienza di questa consonante potrebbe essere stata indotta dal fatto che la prima parola del training ('pabalava') presentava questa consonante all'inizio della parola e quindi era fondamentale per riconoscere l'inizio della sequenza / parola.

Bisogna prendere il considerazione che il soggetto ha richiesto 17 sessioni di allenamento prima di esperire le parole senza sforzo cognitivo. Certamente il compito richiesto richiede inizialmente uno sforzo attentivo (ci vuole molto tempo prima di passare da una sensazione "cognitiva" a una "percezione").

Si noti che l'andamento del training mostra chiaramente un effetto di apprendimento: quando il soggetto apprende una nuova parola, la parola appresa in precedenza viene individuata un

numero inferiore di volte, per poi ritornare alla performance acquisita dopo qualche blocco di training (si veda tab.)

Si è parlato di 'contesto simil-ecologico' dal momento che un sordo profondo potrebbe utilizzare il dispositivo come convertitore 'online' di un parlante (si pensi ad una chiamata telefonica). In letteratura, non sono presenti studi che mostrano buone prestazioni con la normale velocità del parlato utilizzando solo dispositivi tattili. Dal momento che anche in questo pilota non si è riusciti a raggiungere una performance elevata (per lo meno in relativamente poco tempo di esposizione al linguaggio tattile), si può concludere che il riconoscimento delle parole alla normale velocità del parlato può essere abbastanza difficile o addirittura impossibile, utilizzando un dispositivo tattile, a meno che non si faccia un lungo periodo di training (come per l'apprendimento del codice morse).

2.3.1.2 Pilota con l'utilizzo della modalità tattile e lettura delle labbra.

Come accennato nell'introduzione alla sostituzione sensoriale, in un contesto in cui non si riesce ad ascoltare l'altro, per esempio ad una festa in cui l'altra persona è distante, il guardare le labbra a volte genera l'impressione di 'ascoltare meglio' l'altro. Come detto, la sostituzione sensoriale prevede che il fornire informazioni sensoriali in una modalità diversa da quella preposta a veicolare un tipo di informazione, può portare ad avere la medesima (o similare funzionalmente) informazione.

Si è svolto, quindi, uno studio pilota, in cui vengono presentate solo informazioni visive e tattili, nella speranza che il soggetto possa 'percepire il linguaggio' o 'avere l'impressione di sentire meglio', utilizzando la stimolazione tattile al posto dell'informazione uditiva, completamente assente.

Si è deciso di includere un periodo di training giornaliero per consentire la familiarizzazione con il dispositivo tattile, con la speranza di una generalizzazione legata all'estrazione di caratteristiche salienti generali, anche a parole non addestrate.

In base ai riscontri riportati dalla letteratura sul training presentati nella parte teorica del presente lavoro, si è adottato un tipo di training gerarchico (dalle parole a intere frasi), combinando un training solo tattile con un training tattile e visivo.

Dal momento che la letteratura riporta che l'apprendimento iniziale è più lento (si ricordi il soggetto riportato da Brooks et al, 1985 che ha appreso 2,7 parole all'ora per le prime 150 parole, e poi ha raggiunto un tasso di apprendimento di 3.9 parole all'ora per le altre 99 parole), ci si aspetta che i nostri soggetti imparino a riconoscere almeno 2,7 parole all'ora, ovvero

almeno 11 parole puramente tattili (secondo PL Scilley, 1980) in meno di una settimana (se si considera una sessione da 45 minuti circa al giorno).
Inoltre, dal momento che la letteratura riporta che persone non esperte nella lettura delle labbra (per es. persone rese artificialmente sorde) riescono a capire circa 50 parole solo con la lettura delle labbra e circa 70 parole con l'aggiunta della stimolazione tattile (si veda De Filippo e Scott, 1997), ci si aspetta una performance similare prima del training e di gran lunga superiore dopo il periodo di training.
Prima di svolgere il pilota, sono stati svolti altri piloti per stabilire le frequenze e il numero di canali da usare (si veda la sezione strumenti e stimoli).

Metodo.

Soggetti

Hanno preso parte allo studio pilota 2 soggetti (maschio di 33 e femmina di 29 anni) di madrelingua italiana.

Strumenti

Sono stati utilizzati un comune pc portatile, le stesse cuffie dello studio precedente 'Marshall Monitor', dei tappi in gomma e il dispositivo chiamato *FeelSpeech*. Il dispositivo *Felspeech* presenta 10 canali, uno per ogni dito (matrice 2x4). Questo dispositivo utilizzato la stessa logica di vocoders descritti nella parte teorica, estraendo una gamma specifica di frequenze, ovvero estrae l'ampiezza generale del discorso e la divide in base di frequenza suddivise per tutte le sue celle. (si veda fig. 4 d)).
In uno studio pilota sono stati utilizzati 10 canali (le frequenze erano state suddivise in: 100 - 200 Hz, 200 - 450 Hz, 450 - 650 Hz, 650 - 850 Hz, 850 - 1200 Hz, 1200 -1500 Hz, 1500 - 1900 Hz, 1900 - 2600 Hz, 2600 - 3800 Hz, 3800 - 7000 Hz), ma dal momento che erano difficili da codificare, nel presente studio sono stati ridotti a 5 canali (100-450 Hz, 450-850 Hz, 850-1500 Hz, 1500-2600 Hz, 2600-7000 Hz). Il dispositivo funziona online, ovvero estrae le informazioni direttamente dalla voce dello sperimentatore che parla. Per questo motivo è stata mantenuta sempre la stessa voce (si ricordi che una voce femminile ha frequenze più elevate rispetto alla voce di un maschio).

Stimoli

<u>Stimoli uditivi:</u>per rendere i soggetti artificialmente sordi, oltre ai tappi di gomma viene trasmesso il rumore rosa ad alto volume, tramite le cuffie.

Stimoli tattili: come accennato nella sezione ‘strumenti’, il dispositivo presenta cinque celle, ciascuna veicola una banda di frequenze (ampiezza generale, il numero dei bottoni avvertiti è proporzionale al range di frequenze di ciascun canale). Le 5 cellule sono state associate con il anulare (frequenze 100-450 Hz), medio (frequenze 450-850 Hz) e indice (frequenze 850-1500 Hz) della mano sinistra, e indice (frequenze 1500-2600 Hz) e medio (frequenze 2600-7000 Hz) della mano destra. (come in fig. 4 d)).

Stimoli visivi: lettura delle labbra dello sperimentatore.

Le parole sono state tratte dall'elenco delle parole CNC di Maryland (si veda fig 10 a)). Le frasi per la formazione sono state formate ad-hoc, includendo le parole apprese dell'elenco delle parole CNC di Maryland (si veda fig 10 b)). Il testo utilizzato per il test (secondo la *Tracking Procedure* di De Filippo e Scott , 1997) è tratto dalla fiaba di Biancaneve dei fratelli Grimm (si veda fig. 11).

Procedura

Entrambi i soggetti svolgono un training di sei giorni per circa 40 minuti al giorno, e due fasi di test, prima e dopo il periodo di training.

Ogni sessione di allenamento comprendeva: 1) 10 minuti di apprendimento solo tattile delle parole (il soggetto porge le spalle allo sperimentatore e ripete la parola dopo che lo sperimentatore l’ha pronunciata), 2) 15 min di apprendimento delle medesime parole con il dispositivo e la lettura delle labbra (lo sperimentatore dice la parola e il soggetto la ripete), 3) 15 minuti di comprensione di frasi con dispositivo e lettura delle labbra (lo sperimentatore dice la frase e il soggetto la ripete).

Il test (pre e post training) è stato svolto secondo la *Tracking Procedure* di De Filippo e Scott (1997) ed è stato calcolato il numero di parole al minuto (10 minuti di lettura). Secondo la procedura (si veda la parte teorica), lo sperimentatore leggeva una frase completa e il soggetto doveva ripetere la frase. Nel caso in cui la frase ripetuta non era corretta, lo sperimentatore dava suggerimenti, rileggeva la frase e il soggetto doveva nuovamente ripeterla.

Risultati

Dopo sei giorni di formazione, il Soggetto1 (maschio) ha imparato a riconoscere utilizzando solo il dispositivo 11 parole, con accuratezza superiore al 75%; mentre il Soggetto2 (femmina) ha imparato a riconoscere con lo stesso livello di accuratezza, 9 parole.

Per quanto riguarda il test (dispositivo più lettura delle labbra), in dieci minuti, nel pre-test il il soggetto1 ha compreso 50 parole, mentre il soggetto2 solo 36 parole.

Nel test post training, il Soggetto1 ha compreso 54 parole, mentre il Soggetto2 ha compreso 49 parole. (si veda la fig. 27 a) e b) in appendice).

Discussione e conclusione

Il presente pilota aveva come fine lo studio dell'efficacia dell'informazione fornita dal dispositivo tattile nella comprensione della lettura delle labbra e quindi 'dell'ascoltare meglio' utilizzando la stimolazione tattile.

Per quanto riguarda l'apprendimento delle parole con la sola stimolazione tattile, il Soggetto1 ha confermato la nostra aspettativa (almeno 11 parole), mentre il Soggetto2 ha imparato un numero di parole un pò sotto la nostra stima (2.25 parole all'ora invece di 2.7 parole all'ora).

Il risultato più scoraggiante rispetto alle nostre aspettative iniziali riguarda i risultati ottenuti nel test. Infatti, dal momento che le prestazioni dei nostri soggetti sono vicine alle prestazioni delle persone non esperte con la sola lettura delle labbra (De Filippo e Scott, 1997), possiamo concludere che l'utilizzo del dispositivo non ha incrementato la comprensione del linguaggio. Ne consegue che il miglioramento presente tra il pre e il post test che i soggetti hanno mostrato rappresenta solo la conseguenza del miglioramento delle loro capacità di lettura del labiale.

Si possono ipotizzare due spiegazioni per questi risultati: o il dispositivo non è stato affatto utile per disambiguare le parole che non erano chiare (limite del dispositivo), o i soggetti hanno ignorato il dispositivo (non hanno fatto utilizzo della stimolazione tattile per comprendere la lettura delle labbra). In entrambi i casi, il dispositivo ha fallito nel suo scopo per comprendere il parlato, anche se hanno dimostrato che i soggetti hanno appreso come usarlo dal momento che riuscivano a riconoscere le parole anche solo con il tatto. Se ne deduce che forse i soggetti non hanno integrato le informazioni del dispositivo con la lettura delle labbra, ma sono in grado codificare, e potenzialmente utilizzare le informazioni provenienti dal dispositivo tattile. Ciò che è mancato di certo è la capacità di generalizzazione delle informazioni tattili per le parole nuove (come mostrato dai risultati del pre e post test).

Dobbiamo anche tener conto del fatto che la procedura di monitoraggio del discorso di De Filippo e Scott (1997) potrebbe essere influenzata dal linguaggio di riferimento. Si ricordi, infatti, che entrambi i soggetti sono di madrelingua italiana, ma che tutta la procedura (training e test) sono stati svolti in lingua inglese. Si può speculare quindi, che le prestazioni dei soggetti possano essere differenti per lingue differenti a quella inglese, e che le performance dei nostri soggetti sarebbe stata migliore se la procedura fosse stata svolta in lingua italiana. Per altro, non esistono dati normativi di riferimento della lingua italiana di bravura della lettura delle

labbra utilizzando questa procedura, per cui non sarebbe stato possibile fare un confronto dei dati non esistenti.

Ovviamente, un'altra considerazione che può essere fatta riguarda la quantità di training necessaria affinché i soggetti riescano a generalizzare le informazioni tattili prima di svolgere un test come quello da noi utilizzato e che quindi i soggetti hanno bisogno di un'enorme quantità di training prima di poter avere un vantaggio dall'utilizzo del dispositivo per la comprensione del linguaggio tramite la lettura delle labbra.

2.3.2 'Ascoltare' con il tatto, con parziale (o residua) informazione uditiva.

2.3.2.1 Pilota con audiolibro.

Dal momento che l'ipotesi della sostituzione sensoriale, prevede che l'informazione tattile possa dare l'impressione di 'ascoltare' meglio, si è cercato di verificarne l'efficacia, nonché la presenza di un effetto facilitatorio, con un dispositivo multicanale.

Si è svolto uno studio pilota in cui un soggetto ha provato ad ascoltare un audiolibro con informazione degradata, utilizzando un dispositivo tattile multicanale.

Ci si aspettava che dopo un certo periodo di utilizzo del dispositivo concomitante all'ascolto dell'audiolibro, il soggetto potesse esperire un'impressione percettiva di comprendere meglio l'audiolibro.

Metodo

Soggetti

Ha preso parte allo studio pilota un soggetto maschio (sperimentatore, età di 37 anni) di madrelingua russa.

Strumenti

Sono stati utilizzati un comune pc portatile, le stesse cuffie dello studio precedente 'Marshall Monitor' e il dispositivo chiamato *FeelSpeech*. Il dispositivo *Feelspeech* utilizzato in questo esperimento, estrae e presenta 18 fonemi (l/b/p/dh/v, d/f/m/zh/g, s/z/k/t/th, n/sh/jh), presentati con 4 diverse stimolazioni tattili (movimento *microswiping* verso sinistra, destra, in alto e in basso) dei bottoni delle prime e ultime file (si veda fig.). Per un limite tecnico, più fomeni sono stati associati ad ogni direzione. In base alla matrice di confusione di Miller and Nicely (1955) (fig. 3 in appendice) i fonemi sono stati distinti in modo tale che le consonanti confondibili presentassero pattern tattile opposto (si veda la fig.12). Visivamente, il dispositivo è come quello utilizzato nell'esperimento psicofisico della discriminazione delle sillabe /BA/ e/VA/ (si veda fig. 4 b) e c)).

Stimoli

Stimoli uditivi: è stato utilizzato un audiolibro (a scelta tra le favole dei fratelli Grimm, "Colors of space", "Supermind", "The angel of the revolution"), filtrato con filtro passa-bassa, che rimuove le frequenze sopra i 650 Hz (dal momento che nei primi stadi di perdita dell'udito vengono perse le alte frequenze).

Stimoli tattili: dispositivo tattile descritto nella sezione 'strumenti'.

Procedura

Giornalmente (un'ora al giorno), per un periodo complessivo di 3 mesi, il soggetto decide la favola che desidera ascoltare, tenendo in mano il dispositivo. Viene chiesto al soggetto di riferire un'eventuale impressione di 'ascoltare' meglio l'audiolibro.

Risultati e discussione

Nello studio pilota, ci si aspettava che dopo un certo periodo di ascolto passivo, il dispositivo influenzasse la percezione uditiva dell'audiolibro con informazione audio degradata. Il soggetto non ha riportato nessuna sensazione di comprensione migliore dell'audiolibro utilizzando il dispositivo. Si è ipotizzato che probabilmente l'ascolto passivo, o la completa assenza di un periodo di training per un dispositivo così complesso, possano aver determinato il mancato riscontro di un beneficio dell'utilizzo del dispositivo.

2.3.2.2 Esperimento con l'audiolibro.

Come accennato, seguendo l'obiettivo di 'ascoltare' meglio utilizzando l'informazione tattile al posto di quella uditiva, si è voluto verificare questo miglioramento utilizzando stimoli uditivi parzialmente degradati (un pò come le prime fasi di perdita dell'udito).

In uno studio pilota, gli sperimentatori hanno ascoltato un audiolibro con informazione acustica degradata, per pochi minuti con e senza dispositivo. Dopo qualche minuto, la sensazione di tutti e tre gli sperimentatori è stata di un 'aumento' del volume dell'altoparlante e la sensazione di sentire meglio il discorso, quando veniva appoggiato il pollice sul dispositivo. Come accennato, solitamente l'utilizzo efficace di questi dispositivi implica molte sessioni di addestramento prima di raggiungere buoni risultati (incremento del 20% di parole al minuto) e la maggior parte dei training a livello delle parole non implicano un aiuto efficace nella comprensione del discorso, ragion per cui, data sensazione immediata di comprensione del linguaggio è stata incoraggiante per la creazione di un esperimento. Si è deciso di confrontare

la comprensione delle frasi con e senza dispositivo, senza svolgere alcun periodo di training. Si è cercato quindi di verificare in modo quantitativo la veridicità o meno della "sensazione" di ascoltare meglio un discorso in corso (riportata qualitativamente).
Ci si aspetta che la percentuale di parole corrette riportate sia più elevata per le frasi in cui i soggetti utilizzano il dispositivo, piuttosto che per le frasi in cui non utilizzano il dispositivo.
Prima di eseguire l'esperimento è stato svolto un pilota per stabilire la lunghezza delle frasi.

Metodo.

Soggetti

Hanno preso parte all'esperimento 20 soggetti normoudenti (16 femmine, 4 maschi, età media). I soggetti sono di madrelingua diverse, ma tutti usano la lingua inglese nella loro vita quotidiana. Tutti i soggetti hanno firmato la dichiarazione di consenso informato e hanno dichiarato di essere in buona salute e di non avere problemi di udito. I soggetti sono stati pagati 10 euro l'ora (*ERC PoC Grant FeelSpeech*). (si veda fig. 4 b)).

Strumenti

Sono stati utilizzati un comune pc portatile, le stesse cuffie dello studio precedente 'Marshall Monitor' e il dispositivo chiamato *FeelSpeech*. Il dispositivo *Feelspeech* utilizzato in questo esperimento, estrae e presenta l'ampiezza generale del discorso. Il numero del bottoni alzati l'ampiezza generale del discorso. In base a 4 range di frequenze (0-650 Hz, 650-1400 Hz, 1400-3000 Hz, 3000-5000 Hz, 5000-6500 Hz), il numero dei bottoni (e quindi la porzione di pelle del pollice stimolata) aumenta (si veda fig.13).

Stimoli

<u>Stimoli uditivi:</u> sono presentate 100 frasi tratte da un audiolibro delle favole di Grimm (tra cui la già citata Biancaneve), filtrato con filtro passa-bassa, che rimuove le frequenze sopra i 650 Hz (dal momento che nei primi stadi di perdita dell'udito vengono perse le alte frequenze). Ogni frase conteneva circa 10 parole (stabilite dopo il pilota per non creare un sovraccarico di memoria di lavoro). (si vedano le frasi in fig. 14).
<u>Stimoli tattili:</u> dispositivo tattile a singolo canale (si veda strumenti). La stimolazione era proporzionale (il numero dei bottoni avvertiti, che seguono una forma a spirale) all'ampiezza presente nel segnale audio degradato. (si veda fig.13)

Procedura

L'esperimento dura circa un'ora. Ogni soggetto svolge una fase di familiarizzazione e la fase di test.

Nella fase di familiarizzazione, i partecipanti ascoltano per 10 minuti, una traccia di audiolibro degradato (differente rispetto a quella che viene presentata per il test).

Nel test, viene presentata una frase per volta dell'audiolibro con audio degradato e viene chiesto ad ogni partecipante di digitare al computer tutte le parole che ha colto dell frase che ha ascoltato. Delle 100 frasi totali, sono presentate metà frasi con il dispositivo e metà senza, in modo random tra i soggetti (metà soggetti iniziano con il dispositivo, metà senza). Nel complesso, quindi, le frasi sono state divise in quattro condizioni, controbilanciate tra i soggetti (risultate dalla manipolazione delle metà frasi presentate con e senza dispositivo, e presentate al primo o secondo blocco). In dettaglio: 1) nella condizione A il dispositivo tattile è in uso al primo blocco con la prima metà delle frasi (secondo blocco senza dispositivo e con la seconda metà di frasi), 2) nella condizione B il dispositivo tattile è in uso al primo blocco con la seconda metà delle frasi (secondo blocco senza dispositivo e con la prima metà di frasi), 3) nella condizione C il dispositivo tattile è in uso al secondo blocco con la prima metà delle frasi (secondo blocco con il dispositivo e con la seconda metà di frasi), 4) nella condizione D il dispositivo tattile è in uso al secondo blocco con la seconda metà delle frasi (secondo blocco con il dispositivo e con la prima metà di frasi).

Risultati

E' stata confrontata la quantità di parole correttamente digitate nelle diverse condizioni. Il t-test appaiato ha mostrato un effetto principale per i blocchi ($p = .00005$). Il secondo blocco è risultato sempre più facile del primo, indipendentemente dall'ordine di presentazione e dall'utilizzo di un dispositivo tattile o meno.

Non è stato riscontrato nessun effetto significativo nel confronto tra le prove con e senza dispositivo in generale ($p= .695$). Nel dettaglio, non sono significative le differenze tra l'utilizzo del dispositivo o meno per la condizione A ($p=.037$), condizione B ($p=.042$) condizione C ($p=.021$) e condizione D ($p=.031$). (si veda fig.28).

Discussione e conclusione

Lo studio svolto mirava a verificare l'esistenza di un effetto facilitatorio della stimolazione tattile sulla comprensione uditiva, in presenza di stimoli uditivi parzialmente comprensibili (così come avviene nelle prime fasi della perdita dell'udito). Data l'impressione iniziale del dispositivo di incrementare il volume e di comprendere meglio il linguaggio, ha dato speranza

per un utilizzo del dispositivo che non prevedesse un periodo di training, risolvendo uno dei problemi maggiori dell'abbandono di questi dispositivi, ovvero la richiesta di un grosso sforzo cognitivo iniziale per il loro utilizzo.

Il mancato riscontro di un vantaggio quantitativo per la comprensione del linguaggio utilizzando il dispositivo, ovvero l'assenza di un incremento significativo del numero di parole correttamente comprese nella condizione con il dispositivo tattile, non ha confermato la nostra sensazione percettiva di comprendere meglio il discorso utilizzando la stimolazione tattile da noi fornita, sconfermando la nostra ipotesi iniziale. Probabilmente la sensazione immediata di maggiore comprensione del linguaggio sia stata dovuta al fatto che questo dispositivo, veicolando l'ampiezza generale del linguaggio, in qualche modo ricorda i dispositivi a canale singolo, che come detto, sono molto utili per seguire il ritmo e la prosodia del linguaggio, che possono dare una sensazione immediata di 'aumento' della comprensione linguistica.

Ovviamente, dal momento che non è stato eseguito alcun periodo di training, ci si attendeva un effetto piccolo, che tuttavia non è stato riscontrato. Si può quindi, ipotizzare che aggiungendo un certo periodo di training (o una familiarizzazione più lunga), si potrebbero ottenere risultati differenti. Inoltre, il nostro risultato mostra che i due blocchi non erano ben equilibrati (un blocco è stato più facile dell'altro), ragion per cui bisogna anche ipotizzare che probabilmente questo tipo di test è molto sensibile alla quantità di parole e al loro utilizzo nel discorso comune e andrebbe bilanciato meglio. Si consideri inoltre, che la performance potrebbe essere molto diversa in presenza di audiolibro della stessa madrelingua dei soggetti, ma ciò non garantisce comunque una differenza tra l'utilizzo del dispositivo o meno. Per esempio, si può ipotizzare che i soggetti non madrelingua possano compiere degli errori di scrittura delle parole comprese, che non vengono rilevate dal programma e quindi non incluse, nell'analisi.

Conclusione sulla sostituzione sensoriale: 'ascoltare' con il tatto.

In questa sezione sperimentale si è cercato di trovare evidenze per una sostituzione sensoriale audio-tattile. In particolare, si è ipotizzato che fosse possibile riscontrare un effetto percettivo di 'ascoltare meglio' un discorso (in completa o parziale assenza di informazione uditiva), utilizzando l'informazione tattile fornita dalle diverse versioni del dispositivo denominato per l'appunto, *Feelspeech*.

Come si è visto dalla rassegna bibliografica, gli studi precedenti con dispositivi tattili hanno riportato periodi di training molto lunghi, ma nessuno di essi supponeva o voleva cercare evidenze in merito alla possibilità di sostituzione sensoriale audio-tattile. La maggior parte di questi studi, infatti, mirava alla commerciabilità dei dispositivi tattili. Si è visto che il training

su singole parole non implicava necessariamente un'efficacia del dispositivo tattile per stimoli presentati in un discorso condotto a velocità normale (perfino nella procedura di monitoraggio, lo sperimentatore è costretto ad esporre frase per frase ed intervenire). In questa sede, quindi, si è cercato non solo di testare la possibilità di comprendere il linguaggio a velocità normale utilizzando il dispositivo, si è cercato anche di ridurre i periodi di training da un anno (Brooks et al, 1985) a pochi mesi (pilota con l'audiolibro).

Gli esperimenti e gli studi pilota condotti in questa sede, manipolando sia l'informazione uditiva (assenza o parziale presenza) che tattile (utilizzo vs. non-utilizzo), che il training (con e senza) non hanno portato a corroborare l'ipotesi di sostituzione sensoriale audio-tattile. In altre parole, l'effetto di facilitazione dell'informazione tattile sulla percezione uditiva non è stato riscontrato, né a livello qualitativo, puramente percettivo, né a livello quantitativo, tramite i test che sono stati utilizzati.

Pertanto, ad oggi, non esistono evidenze sulla possibilità di una effettiva sostituzione sensoriale audio-tattile, nonchè la sensazione di poter 'ascoltare' tramite il tatto.

2.4 Rappresentazione della realtà: sopramodale o modalità specifica.

2.4.1 Studio con mappe.

Nella parte introduttiva del presente lavoro, sono stati esposti i problemi concernenti la rappresentazione mentale della realtà. Come detto, questa espressione indica il 'ripresentare' o 'rendere (di nuovo) presente' la realtà che ci circonda, o ad una traduzione in simboli (degli oggetti della percezione), o produzione di altri simboli, che 'rappresentano' qualche cosa (Craik, 1943). Si è fatto cenno, inoltre, alle ipotesi sui processi e fattori che contribuiscono alla creazione di una rappresentazione, facendo riferimento, in particolare, alla discussione circa l'esistenza di una equivalenza funzionale tra le diverse modalità sensoriali, ovvero la presenza di analogie strutturali intrinseche che vengono veicolate dai sensi e che portano alla medesima rappresentazione mentale degli oggetti o eventi.

In questo lavoro si è voluto contribuire alla verifica dell'ipotesi dell'equivalenza funzionale delle modalità sensoriali, assunzione, per altro, fondamentale per l'ipotesi della sostituzione sensoriale. In particolare è stata presa in analisi la rappresentazione spaziale di mappe. Nella parte introduttiva sono state analizzate alcune delle componenti e i processi implicati nella rappresentazione spaziale degli ambienti circostanti. Come già detto questi studi hanno dimostrato che i soggetti ciechi (sia congeniti che acquisiti) sono in grado di riprodurre con alto livello di precisione delle mappe composte da suoni, e di identificare correttamente la mappa target in mezzo ad altri distrattori (si veda Olivetti et al., 2007 e Olivetti, 2011). Questi dati sono stati interpretati a favore della teoria di una rappresentazione sovramodale dello spazio per mappe. Per confermare questa teoria, si è pensato di svolgere uno studio con normovedenti bendati, usando una procedura inversa rispetto a quella utilizzata con i ciechi. Infatti, mentre i ciechi partivano da un'esplorazione poco familiare (quella uditiva) e procedevano ad un riconoscimento con una modalità sensoriale da loro molto praticata (quella tattile), si è pensato che fosse opportuno, per confermare questa ipotesi, far partire i partecipanti normovedenti bendati da un'esplorazione poco familiare (quella tattile), per poi riconoscere tra distrattori di mappe sonificate, la mappa esplorata. In questo modo, se indipendentemente dal tipo di modalità si riconosce correttamente la mappa target, si avrebbe una prova ulteriore in favore di una rappresentazione amodale delle mappe.

Metodo

Soggetti

28 partecipanti normovedenti bendati (14 femmine, 14 maschi, età media 23,93, deviazione standard 2,97) hanno preso parte all'esperimento volontariamente. I partecipanti sono studenti di diverse facoltà (psicologia, ingegneri, scienze umane, arte e medicina). Tutti i partecipanti hanno dichiarato di essere in buona salute e di non aver avuto nessuna diagnosi psichiatrica. Le capacità motorie e uditive erano normali in tutti i soggetti. Ciascun partecipante ha firmato il consenso informato prima di svolgere l'esperimento. Prima di cominciare veniva detto ai partecipanti che stavano per svolgere un esperimento con stimoli tattili e uditivi e lo scopo della ricerca era quello di indagare la percezione cross-modale, mentre non veniva detto loro che tra le mappe, oltre ai distrattori, era presente anche la mappa target.

Stimoli

Sono state utilizzate una mappa tattile e tre mappe sonificate.

La mappa tattile è una mappa di plastica termoformata (ho scelto la mappa termoformata, chiamata MAP2, che rispetto alle tre diverse mappe utilizzate precedentemente presenta complessità media, si veda Delogu et al. 2006 e 2010). La mappa, grande 42 cm x 30 cm presenta la forma degli Stati Uniti. Ogni stato interno, presenta una tra tre diverse trame, ad indicare un diverso tasso di disoccupazione, E' stata utilizzata la trama a righe per il tasso di disoccupazione basso, la struttura a spina di pesce per il tasso di disoccupazione medio e la struttura punteggiata per il tasso di disoccupazione alto (come in fig. 15).

Le mappe sonificate sono state create tramite il sistema di sonificazione chiamato iSonic (Zhao et al., 2008). Come citato sono state utilizzate tre mappe sonificate: la mappa target (identica alla mappa termoformata) e due mappe distrattori, diverse in termini di carico di memoria, una più semplice (chiamata MAP1) e una più complessa (chiamata MAP3). Le mappe distrattore sono identiche alla mappa target per la forma e variano solo nella distribuzione del tasso di disoccupazione: la mappa più semplice presenta una chiara distinzione delle tre aree dei livelli di disoccupazione, mentre la mappa più complessa presenta i tassi di disoccupazione distribuiti in modo più variegato e realistico. Per le mappe sonificate sono stati usati tre diversi suoni associati ai diversi livelli di disoccupazione: suono acuto per l'alto tasso di disoccupazione (corrispondente a un sol, a 783,9 Hz) , un suono intermedio associato al tasso medio (corrispondente a un a 659,2 Hz) e un suono grave per il tasso basso (corrispondente a un do a 523,2 Hz). (si veda fig. 15 bis in appendice).

Inoltre, erano presenti altre indicazioni in base al tipo di esplorazione. Nell'esplorazione della mappa sonificata con il touchscreen (si veda procedura) il passaggio dal confine degli stati uniti con il 'nulla' veniva indicato da un suono simile a uno *switcht* di chitarra. Mentre, nell'esplorazione della mappa sonificata con la tastiera (si veda procedura), la fine del confine veniva indicato con una voce che indicava il nome del confine (per esempio "confine sinistro") (come Delogu et al., 2010).

Procedura

Come detto, i partecipanti sono stati bendati. Prima dell'esperimento lo sperimentatore ha chiesto ai partecipanti di esplorare la mappa tattile, capire come era fatta e di memorizzarla il meglio possibile. Veniva spiegata quindi l'associazione delle varie tessiture con il corrispettivo tasso di disoccupazione associato. Come già detto, tutti i partecipanti hanno esplorato attivamente la mappa tattile chiamata MAP2 (si veda la sezione stimoli). Sono stati dati circa 5-10 minuti, anche se i partecipanti erano liberi di decidere la durata effettiva dell'esplorazione e dire quando si sentivano sicuri di aver capito come era fatta. (si veda fig. 19).

Dopo l'esplorazione tattile, lo sperimentatore verificava le abilità di localizzazione e discriminazione di suoni, schioccando le dita a destra e a sinistra della testa del soggetto, e facendo ascoltare i tre suoni presenti nella mappa. Spiegava inoltre che di seguito avrebbe fatto esplorare delle mappe sonificate (spiega l'associazione dei tre suoni associati con i tre livelli di disoccupazione) e che per ognuna avrebbe chiesto al soggetto di dire quanto sia somigliante con la mappa termoformata esplorata in precedenza (da 0 a 10). L'esplorazione delle mappe, anche essa di 5-10 minuti a mappa, è stata svolta o tramite touchscreen o con la tastiera (metà dei partecipanti ha usato il touchscreen e l'altra metà la tastiera, in modo random). Al termine dell'esperimento, il partecipante veniva sbendato e gli veniva chiesto di indicare (da 0 a 6) quanto fosse stata chiara la rappresentazione della mappa tattile, le mappe sonificate e se la discriminazione dei suoni fosse stata facile.

Risultati

E' stata svolta l'analisi della varianza utilizzando il fattore mappa (target e distrattore) come il fattore indipendente all'interno dei soggetti.

I punteggi dati alla mappa target sono significativamente più alti (media = 6,21, SD = 2,27) rispetto ai punteggi dati alla mappa distrattore più semplice (media = 3,86, SD = 1,96) e alla mappa distrattore più complessa (media = 4,26, SD = 2,02).

E’ stata svolta un’ ANOVA mista utilizzando il fattore mappa (target e distrattore) come fattore all'interno dei soggetti e il fattore modalità di esplorazione (touchscreen e tastiera) come fattore tra soggetti. I risultati mostrano un effetto significativo per il fattore principale mappa ($p = .002$), che conferma il fatto che la mappa target sia stata identificata rispetto ai distrattori. Non è stato trovato alcun effetto di interazione, il che significa che non è stata riscontrata nessuna differenza tra i soggetti che hanno esplorato le mappe sonificate con il touchscreen o la tastiera (7 partecipanti per ciascuna modalità di esplorazione).
Tutti i partecipanti hanno riportato sul questionario di avere un'idea chiara della mappa tattile e delle mappe sonificate e che i suoni erano chiari.

Discussione e conclusione

Come accennato nella parte introduttiva, la rappresentazione spaziale è fondamentale per lo svolgimento di molte delle nostre attività quotidiane. Ad oggi, tra le ipotesi più corroborate sui processi di costruzione della rappresentazione spaziale, la letteratura riporta l’ipotesi sovramodale.
Questo esperimento è stato svolto con lo scopo di verificare se le modalità tattile e uditiva possono fornire una rappresentazione mentale equivalente di mappe, solitamente esplorate e memorizzate visivamente. In altre parole, se possa essere confermata la teoria di rappresentazione spaziale sovramodale, per cui l’equivalenza funzionale delle varie modalità, fa sì che indipendentemente dalla modalità sensoriale utilizzata in fase di esplorazione, si possa avere una rappresentazione spaziale accurata.
I risultati ottenuti mostrano che i soggetti (71% dei partecipanti) hanno riconosciuto la mappa target tra i distrattori o hanno dato punteggi alti, rispetto alle altre mappe.
Una critica che può essere fatta riguarda il fatto che nello studio sono state utilizzate due mappe distrattore che presentavano complessità differente, e di conseguenza, richiedevano anche un carico di memoria differente. Ragion per cui, anche se i soggetti non sapevano a priori della presenza della mappa target tra le mappe sonificate (veniva chiesto solo di dare un punteggio di somiglianza alle mappe sonificate), le risposte dei partecipanti possono riflettere la differenza della complessità della mappa target rispetto ai distrattori.
Ad ogni modo, i risultati sono incoraggianti per la teoria della rappresentazione sovramodale spaziale. Sebbene infatti, modalità sensoriali veicolano proprietà sensoriali specifiche, è probabile che vi siano anche delle analogie strutturali che permettano un’equivalenza funzionale per la costruzione di una rappresentazione spaziale comune, definibile come ‘crossmodale’.

Ovviamente, ulteriori studi sono necessari per confermare questa interpretazione dei risultati, in primis, proprio rispondendo all'accusa plausibile sui livelli di difficoltà delle mappe, variandone la complessità.

Un'eventuale conferma della teoria sovramodale della rappresentazione spaziale potrebbe essere in linea con l'ipotesi della sostituzione sensoriale.

3. DISCUSSIONE

Il presente lavoro ha avuto come finalità lo studio di alcuni fenomeni completamente o parzialmente non indagati, tra cui: le finestre temporali di integrazione audio-tattile, fenomeno del binding audio-tattile di stimoli linguistici e le ipotesi di sostituzione sensoriale e della rappresentazione sovramodale della realtà. La parte sperimentale del presente lavoro, ha cercato quindi di fornire dei contributi empirici in merito.

La letteratura sul binding audio-tattile ha mostrato delle lacune in merito allo studio delle finestre temporali audio-tattili per stimoli presentati in flussi ritmici. Si è svolto quindi, uno studio per valutare la capacità di detezione di stimoli asincroni, in condizione unisensoriale (solo tattile e solo uditiva) e multisensoriale (audio-tattile), presentando i flussi ritmici in modo sincrono o distanziati per distinguere un compito di tipo percettivo da uno mnestico-immaginativo. I risultati hanno mostrato che la detezione delle asincronie è stata molto difficile in tutti i compiti in cui i flussi sono stati distanziati temporalmente, mentre la detezione di tutte le asincronie è stata semplice nella condizione sincrona, solo per il compito unisensoriale uditivo. Questo dato corrobora i precedenti risultati in letteratura circa la velocità di elaborazione uditiva (REF) di gran lunga più veloce rispetto a quella tattile (REF). Il nuovo dato, riguarda quindi la detezione delle incongruenze audio-tattili per flussi, che risulta avere prestazione 'intermedia' tra le condizioni unisensoriali. Si può quindi ipotizzare, che a livello temporale la detezione audio-tattile richieda velocità intermedia rispetto alle due elaborazioni unisensoriali. D'altro canto, bisogna anche notare che nella condizione audio-tattile, la presenza di asincronie durante il test ha fatto sì che non si verificasse un'integrazione audio-tattile, ma è viene mantenuta una percezione degli stimoli uditivi e tattili come distinti.

Per quanto riguarda la ricerca di evidenze sperimentali dell'esistenza dell'integrazione audio-tattile con stimoli linguistici, sono stati svolti uno studio completo (esperimento psicofisico di discriminazione delle sillabe /BA/-/VA/ e uno studio pilota (ricerca di un corrispettivo effetto McGurk audio-visivo con le modalità audio-tattile). È stato svolto inoltre, uno studio pilota con toni musicali. Riguardo allo studio di discriminazione delle sillabe /BA/-/VA/ il fatto che i soggetti abbiano riportato risposte contrastanti all'effettiva stimolazione uditiva solo con stimolazione tattile sopra soglia (blocco 'biased strong'), come già detto, riporta al dibattito e problema, circa l'interpretazione dei risultati. In altre parole, non è

possibile distinguere se ciò che è stato da noi riscontrato rifletta l'effettivo cambiamento nella percezione, in modo genuino, o un cambiamento di criterio nella scelta, ovvero nel rispondere seguendo la modalità tattile (sempre disponibile) rispetto a quella uditiva (incerta). Il problema dell'interpretazione dei dati si ripropone in vari lavori in ambito crossmodale svolti precedentemente. Si prenda ad esempio la "*parchment skin illusion*" di cui si è già parlato. In questa illusione i partecipanti devono dare un giudizio di ruvidità e cercare di ignorare la concomitante stimolazione acustica (ma questa contiene un'informazione semanticamente associabile a quella tattile). Dal momento che l'informazione acustica si deve ignorare coscientemente (viene richiesto ai partecipanti), ma può influenzare la decisione o i meccanismi che portano alla risposta, è difficile poter dire con certezza da cosa dipenda la percezione finale.

Chiarire e indagare il contributo dell'aspetto sensoriale versus i processi di elaborazione decisionale nell'integrazione multisensoriale è cruciale per accertare la natura dei meccanismi sottostanti (Soto-Faraco e Deco, 2009). Alcuni tentativi che sono stati fatti in passato per superare questi limiti hanno usato stimoli distrattori che non variano lungo la dimensione misurata nella modalità target. Un approccio svolto è che lo stimolo distrattore è neutrale rispetto alla dimensione in cui deve essere svolta la decisione e rimane invariato durante tutto l'esperimento. Se il distrattore anche in questo caso produce un effetto nei giudizi dati sulla modalità rilevante, allora si assume che l'effetto non è dovuto ad una influenza allo stato decisionale (es. REF). Un altro approccio, è che lo stimolo distrattore varia in modo ortogonale al set di risposte che sono utilizzabili dal partecipante (si vedano REF). Un'altra strategia che può essere utilizzata è quella di usare modelli psicofisici che permettono di discriminare il contributo della sensibilità percettiva dagli effetti puramente decisionali, come per esempio negli studi della teoria della detezione del segnale (REF). Infine, un altro modo per isolare la componente percettiva delle interazioni 'inter-sensoriali' è di utilizzare un distrattore che sia sufficientemente intenso per influenzare l'elaborazione percettiva ma debole abbastanza per essere non identificabile dal partecipante. Se infatti, non c'è una percezione cosciente del distrattore, allora l'informazione che veicola si assume che non dovrebbe influenzare la decisione ad un livello cosciente (per esempio REF).

Nello studio psicofisico svolto, si sono tenute in conto queste considerazioni, per cui si è cercato di confrontare le condizioni in cui entrambe le modalità fossero ambigue vs. la condizione in cui una modalità sensoriale è reperibile e l'altra rimane ambigua (strategia del distrattore neutro). Dal momento che i nostri risultati mostrano che la modalità tattile influenza la risposta di decisione della sillaba, solo quando è sopra soglia, e non quando è sotto soglia,

non possiamo rispondere con certezza circa il fatto che la risposta data dai soggetti sia legata ad un fenomeno puramente percettivo, oppure ad una strategia di risposta.
Nel nostro esperimento, il cambiamento di criterio nel dare la risposta potrebbe essere in parte una conseguenza delle aspettative legate ai biases che sono stati inavvertitamente introdotti dal metodo (premono più volte BA e quindi 'sostano' sul tasto di risposta 'BA') o sotto forma degli effetti della compatibilità stimolo-risposta tra l'informazione veicolata dalla modalità irrilevante (es. tattile) e la risposta richiesta dalla modalità target (uditiva) (REF).
Una considerazione che può essere fatta, riguarda anche la chiarezza/ambiguità dell'informazione veicolata da entrambe le modalità. Per esempio, una delle interpretazioni che viene data in merito al fenomeno del McGurk originale (conversazione con Daniel Penitzer), è che la sillaba /BA/ sia acusticamente debole o ambigua, e che la sillaba /GA/ sia visivamente ambigua, ragion per cui le due modalità vengono ad integrarsi e sono 'forzate' ad integrarsi perché non completamente accessibili.
Una critica che può essere metodologicamente, riguarda il tipo di compito richiesto. Solitamente infatti, negli articoli che presentano un'influenza crossmodale con una stimolazione sottosoglia, il compito richiesto non è di tipo discriminativo, come da noi richiesto, ma di giudizio percettivo. Esempi sono l'effetto McGurk audio-visivo (McGurk H, MacDonald J (1976), l'effetto ventriloquo (Alais e Burr, 2004), il "bounce/pass" effect (Sekuler R, Sekuler AB, Lau R, 1997), studi sulla numerosità di flash e bips (Shams, e coll, 2000) e di movimento apparente in base all'altezza di suoni (Maeda e coll., 2004). Si noti però, che anche le risposte soggettive riportate dai soggetti non potevano essere considerate come rilevanti nel giudizio di integrazione audio-tattile, dal momento che, come già detto, si è verificato un fenomeno di adattamento uditivo, legato alla ripetizione degli stimoli uditivi e totalmente indipendente dalla compresente stimolazione tattile.

Un'altra considerazione che è stata fatta riguarda i tempi di apprendimento delle associazioni audio-tattili. Visto che, di fatto, si è scelto di usare una stimolazione tattile non naturale, ma arbitraria, ci si è chiesti se un maggiore apprendimento (si vedano i tempi di apprendimento impiegati durante il pilota del McGurk audio-tattile) poteva portare a corroborare l'integrazione audio-tattile pura, per stimoli linguistici. Il problema dei tempi di apprendimento di un'associazione, o semplicemente, i tempi richiesti affinché si possa verificare una integrazione tra due modalità sensoriali, rimane un argomento ancora aperto, e accomuna sia l'esperimento psicofisico di discriminazione di sillabe, che il pilota del simil McGurk audio-visivo. La letteratura riporta che nel caso dell'apprendimento percettivo, che comporta un miglioramento della discriminazione tra stimoli, può avvenire molto rapidamente,

dopo alcune prove, e a volte può richiedere molto tempo, ore (REF) o giorni (REF). In extremis, si potrebbe perfino pensare che vi siano finestre temporali di apprendimento così piccole che è difficile misurarle (Jain et al. 2010). Si può immaginare un continuum di apprendimento associativo (tra due segnali, in questo caso due informazioni provenienti da modalità sensoriali differenti), in termini di assolutamente facile e immediato, a completamente non apprendibile (REF). Alla domanda 'cosa può essere associato?' Jain et al (2010) affermano che l'apprendimento di associazioni arbitrarie per un sistema percettivo sarebbe inefficiente. Ragion per cui la capacità di monitorare continuamente le correlazioni per un largo set di appaiamenti segnale/attributo, che alcune associazioni che non risultano essere legate da una relazione causale vengono trattate dal sistema come accidentali (il che ricorda i diversi modelli di apprendimento bayesiano e di chance).

Secondo Backus e colleghi (REF) è necessario un periodo di apprendimento lungo tutta la vita, per poter estrarre degli elementi intrinseci. Gli autori affermano che il fatto che l'apprendimento avvenga o meno può essere interpretato come conseguenza se un sistema crede implicitamente che il segnale, a priori, è informativo sulle proprietà stimate dell'ambiente. In questa ottica, forse l'informazione che è stata veicolata non è stata giudicata dal sistema come sufficientemente importante o affidabile per essere appresa. L'apprendimento di stimoli che hanno scarsa informazione (ovvero, scarsamente disambiguati) può risultare più forte (REF), e può rendere possibile il reclutamento di segnali quali la tessitura (REF), disparità verticale (REF) e la forma (REF), non reclutati in caso di stimoli fortemente disambiguati. (Jain et al., 2014). Secondo Backus e colleghi (si veda Jain, Fuller & Backus, 2014) tra i fattori che favoriscono (necessari ma non sufficienti) l'integrazione percettiva ci sono l'intrinsecità (ovvero l'esistenza di fattori che sono intrinseci al sistema percettivo), l'associazione degli stimoli per un periodo di tutta la vita, la simultaneità e la scarsità dell'informazione presente nello stimolo. Per segnali intrinseci si intendono i segnali veicolati dagli stessi elementi mostrati come lo stesso oggetto, mentre per segnali estrinseci si intendono segnali che non sono connessi con l'oggetto la cui apparenza è stata insegnata. Gli autori affermano che l'assenza di apprendimento di un'associazione indica che il sistema percettivo implicitamente crede che il segnale non è informativo, o quantomeno non dovrebbe essere utilizzato per trarre informazioni in merito alle proprietà delle scene rappresentate dal percetto. Continuano affermando che questo meccanismo serve a prevenire l'apprendimento sporadico dovuto a correlazioni fortuite presenti nell'ambiente, a costo di perdere anche un segnale utile. Secondo questo pensiero, quindi, l'apprendimento dovrebbe verificarsi nei casi in cui segnali aggiuntivi sarebbero utili per disambiguare, così come accade nei training con stimoli con scarsa informazione. In questi

casi, infatti, l'apprendimento di stimoli con scarsa informazione superano questa riluttanza del sistema di reclutare un segnale estrinseco. In modo similare, Michel and Jacobs (2007) suggeriscono che l'apprendimento è relativamente facile quando consiste nel modificare una relazione già esistente (parametro di apprendimento), mentre apprendere nuove relazioni (apprendimento strutturale) risulta essere difficile o impossibile.

Un'altra considerazione che può essere fatta riguarda il 'contenuto' delle frequenze (REF). Ci si può chiedere quali siano le interazioni percettive differenti rispetto ad effetti quali l'effetto McGurk, in merito specificatamente alla tipologia di segnale uditivo. Una possibilità è che in questi esempi le correlazioni audio-visive sono frequentemente riscontrate in ambienti naturali e siamo esposti ad essi a partire dall'infanzia, oppure potremmo ipotizzare che i nostri sistemi percettivi siano predisposti ad accettarli perché sono comuni quotidianamente. Probabilmente il linguaggio presenta delle differenze in questi termini. Sebbene infatti, abbiamo utilizzato delle sillabe, per certi versi forse non sono riconducibili a elementi semplici, o che il linguaggio richieda informazioni extra (per esempio, una stimolazione tattile più complessa) per essere attribuibile ed associabile con la stessa sorgente. La domanda che consegue, riguarda quindi, quale sia il ruolo della semantica dell'informazione veicolata da entrambe le modalità nei processi di integrazione.

Purtroppo, anche lo studio pilota con i toni musicali non ha rivelato un effetto di integrazione audio-tattile a livello empirico (mentre a livello percettivo il soggetto talvolta ha percepito dei toni intermedi nelle prove incongruenti, si veda discussione in merito).

Secondo Ernst (2007) questo riporta al 'problema della corrispondenza', per cui il nostro cervello deve capire quali segnali sensoriali corrispondono ad un altro per sapere quali segnali combinare opportunamente. Egli ritiene che si può creare un 'rimappaggio', ovvero il cervello può imparare a considerare determinati stimoli insieme ed integrarli, oppure lasciarli separati. In altre parole, l'autore afferma che è possibile formare una nuova mappature, ovvero corrispondenze uniche (o correlazione statistica) tra segnali diversi se una variabile è informativa sull'altra (ovvero c'è una ridondanza) allora è ipotizzabile che il cervello integri i due segnali. Egli ipotizza inoltre, che il nostro cervello deve poter imparare e integra segnali nuovi a cui non è stato esposto prima (nel nostro caso stimoli linguistici/toni musicali con informazione tattile arbitraria completamente nuova) per adattarsi all'ambiente. In sintesi, l'esposizione a nuove correlazioni tra segnali sensoriali può portare a un segnale ad avere nuovi effetti percettivi (risultato della loro integrazione) (Di Luca et al. 2010) ed è un processo adattativo (Ernst, 2007).

Per quanto riguarda la ricerca di evidenze a supporto della teoria della sostituzione sensoriale, due studi pilota, in cui si è cercato di verificare se la modalità tattile può veicolare informazioni uditive che mancano del tutto, ed un esperimento e uno studio pilota, in cui la stimolazione tattile veicola una parte dell'informazione uditiva degradata (comprensione di un audiolibro con traccia non chiara).

I due studi pilota in assenza di informazione uditiva, sono stati svolti per valutare meglio la capacità di apprendimento solo tattile (primo pilota) e tattile più lettura delle labbra (secondo pilota). Si è valutata quindi, la capacità di apprendimento in entrambi i casi e se l'utilizzo dello strumento desse una sensazione di 'sentire meglio' in assenza di linguaggio, nelle condizioni in cui si è utilizzato il tatto e la lettura delle labbra.

Nel primo studio pilota si è ottenuto un risultato incoraggiante, se si considera la difficoltà del compito, soprattutto per il carico attentivo richiesto. Inoltre, si tenga presente che il soggetto verbalizzava la risposta, il che implica un ulteriore carico cognitivo richiesto dal compito. Ad ogni modo, volendo riuscire a comunicare con un sordo al telefono, l'apprendimento di un intero vocabolario di una lingua richiederebbe un'enorme quantità di training, ragion per cui, al giorno d'oggi, verrebbero preferite altre applicazioni, già presenti nei cellulari (quali, ad esempio, la digitazione automatica sotto dettato di parole). Nel secondo pilota, si è cercato di utilizzare un training che fosse ottimale, rispetto alle conoscenze sul training (si veda la disquisizione sul training e la rassegna di Galvin et al., 2001). Per valutare se l'informazione tattile da noi fornita potesse incrementare la capacità di comprensione linguistica nella condizione in cui era permessa anche la lettura delle labbra, si è utilizzato il test, già citato, del *Tracking method*, di De Filippo e Scott (1978). Come già riportato, i due partecipanti non hanno utilizzato l'informazione tattile nel comprendere il linguaggio, e non hanno riportato una sensazione di 'ascoltare meglio' che fosse possibile indicare come unicamente conseguente alla stimolazione tattile. Il miglioramento riportato, era legato solo ad un apprendimento della lettura delle labbra. La critica che ne consegue, quindi, potrebbe riguardare i tempi di apprendimento. Anche in questo caso, è possibile che per far si che si verifichi un fenomeno di sostituzione sensoriale siano richiesti tempi maggiori di confidenza con lo strumento. Inoltre, come accennato, la metodologia utilizzata non consente di testare un'effettiva sostituzione sensoriale, ma l'eventuale capacità di 'generalizzazione' dell'informazione tattile trasmessa, che non si è riscontrata.

Nel verificare la sostituzione sensoriale in compresenza di informazione uditiva degradata, l'esperimerimento e lo studio pilota hanno due metodologie simili, ma sonos stati svolti con due diverse tipologie di dispositivo. Come già detto, entrambe le versioni dello

strumento di sostituzione sensoriale utilizzato, hanno superato i limiti dei precedenti dispositivi tattili, abbandonati (si veda la descrizione del *FeelSpeech* nel capitolo sulle evidenze sperimentali). Nell'esperimento di comprensione dell'audiolibro con informazione degradata, non si sono riscontrate differenze di prestazione tra le frasi ascoltate con o senza lo strumento di sostituzione sensoriale da noi utilizzato e nello studio pilota il soggetto non ha compreso maggiormente il testo. In altre parole, l'utilizzo del *FeelSpeech* non ha incrementato il livello di comprensione, per cui la stimolazione tattile non è riuscita a veicolare le informazioni mancanti nella traccia dell'audiolibro. Ne consegue che le informazioni veicolate dal *FeelSpeech* non sono state sufficienti per integrare le informazioni mancanti della traccia degradata dell'audiolibro. Anche a livello percettivo, malgrado è stata riportata dai partecipanti la percezione di un 'aumento' di volume, ciò non ha portato a riscontrare un incremento del numero delle parole comprese. Ad oggi, la comprensione del linguaggio rimane un processo per lo più sconosciuto. Studi in merito alla percezione acustica del linguaggio hanno riportato solo alcune difficoltà, per esempio nella discriminazione tra alcune consonanti piuttosto che altra (si veda la già citata 'confusion matrix' di Miller e Nicely, 1955, fig. 3 in appendice) e le principali difficoltà delle persone con deficit uditivi per cui le difficoltà maggiori conseguono alla perdita della percezione delle vocali (si veda la tabella riassuntiva riportata da Galvin et al., 2001). Come accennato, nel nostro esperimento si è cercato di simulare la perdita delle frequenze alte, così come accade nei primi stadi della sordità. Probabilmente, quindi, il dispositivo da noi usato non è riuscito a compensare le frequenze alte, o quantomeno l'informazione veicolata non è stata sufficiente per incrementare la comprensione linguistica. Si deve comunque tenere in considerazione, il fatto che la simulazione di esclusione delle alte frequenze potrebbe non essere completamente indicativa di ciò che accade nella popolazione dei sordi o persone con deficit uditivi (visto che soprattutto nelle fasi iniziali, i deficit uditivi sono altalenanti) e che i nostri partecipanti sono normoudenti, quindi, completamente 'naive' alla condizione proposta dal test.

Si tenga presente che a differenza degli studi esistenti in letteratura, i partecipanti non hanno effettuato nessuna fase di apprendimento, come detto, per verificare un eventuale effetto 'automatico' dello strumento sulla percezione uditiva. E' probabile che essendo il linguaggio un tipo di stimolo complesso, sia necessaria una fase di apprendimento, nonostante la stimolazione tattile utilizzata rispecchi in modo parallelo ('naturale') lo spettro acustico, ovvero sia una stimolazione intrinseca allo stimolo linguistico originario (si veda su la discussione sull'apprendimento). Come si è visto, finora il sistema di sostituzione sensoriale tattile che sembra essere più efficace riguarda il metodo Tadoma, che implica l'utilizzo delle

mani e utilizza la ‘vibrazione’ per decodificare le informazioni linguistiche, il che corrobora l’ipotesi di ‘naturalezza’ e ‘intrinsecità’ richiesta per la comprensione del linguaggio attraverso il tatto. Ne consegue, che lo strumento da noi usato, non presenta una immediata e automatica integrazione delle informazioni veicolate tramite il tatto, con le restanti, disponibili, uditive. Nel nostro ‘azzardo’ a voler verificare una sostituzione sensoriale (seppure non completa) in modo automatico, senza training, conferma l’esigenza di un periodo di training, probabilmente legato ai tempi di plasticità corticale richiesti affinché il processo abbia luogo.

Infine, si è cercato di verificare se la rappresentazione mentale di mappe, esplorate con il tatto e con l’udito, potessero essere funzionalmente equivalenti, nonché si è cercato di fornire un contributo empirico dell’ipotesi sopramodale della rappresentazione della realtà. Il grande limite dell’esperimento che si è svolto è costituito dal fatto che sono state utilizzate mappe distrattori di complessità differente rispetto alla mappa target, ragion per cui andrebbero fatti ulteriori approfondimenti in merito. Il giudizio di similitudine che hanno dato i soggetti, quindi, potrebbe riflettere la diversità di complessità piuttosto che la diversità della mappa target di per sé. Rimane quindi da confermare se tatto e udito forniscono la stessa rappresentazione mentale, o equivalente funzionalmente per svolgere il compito richiesto.

In conclusione, i contributi sperimentali svolti hanno in parte chiarito, ed in parte aperto nuove domande sulle questioni inerenti le finestre temporali di elaborazione percettiva, integrazione, sostituzione sensoriale e rappresentazione nelle modalità tattile e uditiva.

4. CONCLUSIONE

Nel presente lavoro si è cercato di contribuire a rispondere a tre questioni aperte della letteratura. L'analisi della letteratura, infatti, ha mostrato chiare lacune in merito ad alcuni aspetti riguardanti i problemi del binding, l'ipotesi sulla sostituzione sensoriale e l'ipotesi di sovramodalità della rappresentazione della realtà nelle informazioni uditive e tattili. Nonostante udito e tatto condividano sia proprietà di trasmissione fisica delle informazioni in entrata (frequenze vibratorie), sia proprietà di elaborazione dell'informazione (sequenziale), alcuni dei fenomeni percettivi che implicano la loro interazione e mutua influenza sono stati finora poco studiati.

Riguardo alla prima tematica presa in analisi, dalla rassegna della letteratura dei problemi di binding audio-tattile, è emerso che la vicinanza spazio-temporale di stimoli tattili e uditivi comporta una loro integrazione e mutua influenza. Un esempio della loro integrazione è l'esperimento comportamentale di Schürmann e coll., 2004, che mostra che il giudizio del volume di due toni viene influenzato dalla concomitante presenza di una stimolazione tattile vibratoria. Come esempio della loro mutua influenza, si ricordino gli studi sull'attività cerebrale in cui si è rilevata un'attivazione delle aree precoci uditive in presenza della stimolazione tattile (REF) o l'attivazione di aree comuni durante lo svolgimento di specifici compiti (REF). Si è anche visto il caso di un paziente con lesione al nucleo ventricolare talamico, che ha riportato un fenomeno di sinestesia audio-tattile, facendo supporre l'esistenza di una plasticità corticale e l'attivazione di aree tipicamente 'tattili' in seguito a stimolazione uditiva (che 'riflette' ciò che viene riportato dalla paziente come esperienza percettiva) (Ro e coll., 2007).

Da questa analisi bibliografica, è emersa una importante lacuna in merito all'assenza di studi che indagano le finestre temporali entro cui il fenomeno di binding audio-tattile potesse verificarsi, utilizzando stimolazioni audio-tattili in flussi ritmici. Si è svolto, quindi, uno studio in merito in cui sono analizzate le finestre temporali di detezione degli stimoli di flussi unisensoriali (modalità uditiva e tattile) e plurisensoriali (audio-tattile). Dallo studio è emerso che il tempo di detezione di stimoli audio-tattili risulta essere intermedio rispetto alle finestre di detezione degli stimoli unisensoriali (uditivi e tattili). Si può quindi supporre, che il binding audio-tattile di flussi ritmici richieda una finestra temporale 'intermedia' rispetto alla finestra temporale uditiva e tattile (la finestra temporale uditiva risulta di gran lunga più piccola rispetto a quella tattile). Questo dato costituisce il primo passo per la programmazione di studi futuri sul binding audio-tattile di flussi ritmici, ancora non indagato. Il fenomeno di binding audio-tattile per flussi ritmici, potrebbe quindi essere diverso rispetto ai fenomeni di binding audio-

tattile precedentemente studiati che hanno utilizzato stimoli presentati singolarmente o coppie di stimoli.

Un altro aspetto che non era stato ancora indagato dagli studi precedenti sul binding audio-tattile riguarda la presenza dell'integrazione audio-tattile stimoli uditivi linguistici. Sebbene infatti, studi precedenti hanno mostrato chiari effetti di integrazione audio-tattile con toni musicali, non sono state riportate evidenze sul fenomeno utilizzando stimoli linguistici. Anche gli studi precedenti con stimoli linguistici e dispositivi tattili con soggetti con deficit uditivi o sordi, miravano a dimostrare l'efficacia dei dispositivi tattili, piuttosto che dimostrare l'esistenza di un'integrazione audio-tattile delle informazioni uditive residue con le stimolazioni fornite dai dispositivi. In alcuni di questi studi viene utilizzata la parola 'integrazione' audio-tattile, ma in realtà riportano solo che le performance in compiti linguistici sono migliori utilizzando le informazioni provenienti da entrambi i canali sensoriali, rispetto all'utilizzo di un solo canale sensoriale. In altre parole, questi studi riportano il vantaggio delle prove multisensoriali rispetto alle prove unisensoriali, il che non implica necessariamente un'integrazione delle due modalità, ma un vantaggio legato alla doppia 'reperibilità' dell'informazione che serve per il compito (si veda discussione in merito a pag.).

Per verificare l'esistenza dell'integrazione audio-tattile con stimoli uditivi linguistici (come fenomeno puramente percettivo) si sono svolti uno studio psicofisico di discriminazione di sillabe e due studi pilota (studio pilota alla ricerca dell'effetto McGurk audio-tattile e studio pilota con due toni). Come visto, nessuno dei contributi empirici effettuati è riuscito a dimostrare la presenza di un effetto di integrazione audio-tattile. Si è visto infatti, che anche l'unico dato potenzialmente a favore per corroborare questa ipotesi, ovvero la presenza di un'influenza della discriminazione delle sillabe quando la stimolazione tattile è sopra soglia, non può essere considerata una prova di un reale effetto di influenza della stimolazione tattile sulla percezione uditiva (si veda la discussione dell'esperimento). Evidenze dell'integrazione audio-tattile a livello percettivo utilizzando stimoli linguistici, sono ancora da dimostrare. Tale riscontro, sarebbe fondamentale non solo per approfondire e capire meglio i meccanismi sottostanti al fenomeno di integrazione audio-tattile per sé, ma anche l'integrazione crossmodale in generale. Portando in extremis, infatti, risulta poco chiaro se si possa parlare di pura integrazione audio-tattile anche negli esperimenti presenti in letteratura, là dove si era interpretato il risultato come prova di integrazione audio-tattile, il che apre nuove domande sulla metodologia di ricerca in ambito crossmodale (si veda la discussione sull'interpretazione dei dati a pag.). Riprendendo le considerazioni di Ernst sul 'problema della corrispondenza', l'integrazione di segnali nuovi è un processo molto importante per l'adattamento ad un nuovo

ambiente. Se si assume che l'integrazione tra stimoli è un importante processo adattativo, questo costituisce un motivo in più per trovare come studiare il fenomeno con stimoli arbitrari.

L'analisi della letteratura in merito all'integrazione audio-tattile include inevitabilmente il prendere in considerazione la mutua influenza e l'importanza e il ruolo di ciascuna modalità sull'altra nel processo integrativo. In particolare, ci si è focalizzati sull'influenza del tatto sulla percezione uditiva, prendendo come cornice di riferimento la teoria della sostituzione sensoriale. Come già citato, la teoria della sostituzione sensoriale afferma l'esistenza di un'equivalenza funzionale delle modalità sensoriali in virtù della quale è possibile veicolare un'informazione di una modalità sensoriale tramite un'altra modalità sensoriale. Portando quindi, in extremis, l'influenza della modalità tattile sulla percezione uditiva, si è ipotizzato che sia possibile 'ascoltare tramite la pelle', ovvero utilizzare informazioni tattili in presenza di informazione uditiva parziale o assente, e avere quindi l'impressione di 'ascoltare meglio' il parlante utilizzando l'informazione tattile. Gli studi precedenti in merito ai sistemi di sostituzione sensoriale utilizzati per veicolare informazioni linguistiche hanno portato a risultato non soddisfacenti. Tutti i dispositivi tattili precedentemente costruiti, sono stati abbandonati e non sono stati riportati casi in cui i soggetti avessero riportato un'impressione percettiva di 'ascoltare meglio' utilizzando tali dispositivi. Sono quindi stati svolti due studi pilota in assenza di informazione tattile (studio pilota con il linguaggio tattile e studio pilota con informazione tattile e lettura delle labbra) e due studi pilota di presenza di informazione uditiva parziale (studio pilota con audiolibro ed esperimento con l'audiolibro). Come visto, anche se a livello percettivo nello studio con l'audiolibro, inizialmente gli sperimentatori hanno avuto l'impressione di 'sentire' meglio utilizzando questo tipo di dispositivo, il test dell'audiolibro non ha confermato la percezione riportata di 'incremento del volume'. Malgrado i nostri risultati potrebbero essere dovuti alla tipologia di test utilizzati (non sufficientemente sensibili per stimare il fenomeno) o al mancato training (si veda discussione dell'esperimento e sui sistemi di sostituzione sensoriale), l'ipotesi di 'sostituzione sensoriale' rimane ancora da confermare. Rimane quindi, ancora da chiarire se l'ipotesi di equivalenza funzionale, ovvero nella fattispecie, la possibilità di veicolare informazione uditiva linguistica tramite stimolazione tattile, sia fattibile in modo efficace e riscontrabile con evidenze empiriche, oltre che alla sensazione di 'ascoltare meglio' utilizzando l'informazione tattile. Date le evidenze circa la plasticità corticale audio-tattile, probabilmente bisogna concludere che ancora non si è riusciti a costruire un sistema di sostituzione sensoriale (dispositivo tattile) in grado di trovare il giusto 'codice' da veicolare e nel modo più 'naturale' possibile per poter essere appreso facilmente. Qualora l'ipotesi della sostituzione potesse avere

delle conferme, e si riuscisse a trovare il modo di veicolare informazioni acustiche linguistiche attraverso la modalità tattile, si potrebbero creare dei dispositivi che potrebbero aiutare soggetti con deficit uditivi a comprendere il linguaggio anche in situazioni in cui attualmente non è loro consentito (per esempio, al telefono).

Per ultimo, dal momento che entrambe le modalità sensoriali 'acquisiscono' informazioni sul mondo circostante in modo sequenziale, ci si è chiesti se la rappresentazione mentale che ne consegue, potrebbe avere degli aspetti comuni o condivisi (fenomeno per altro poco indagato in letteratura). In altro modo, se è vero che esiste un'equivalenza funzionale, così come è supposto dalla teoria della sostituzione sensoriale, anche la rappresentazione che ne consegue potrebbe essere di tipo 'sopramodale', cioè, tutte le modalità sensoriali dovrebbero veicolare delle informazioni equivalenti funzionalmente che consentono la costruzione di una medesima, o anch'essa equivalente funzionalmente, rappresentazione della realtà. Si è quindi testato se l'esplorazione tattile e riconoscimento uditivo di mappe sonificate portasse a riconoscere una mappa target (esplorata tattilmente) tra varie mappe sonificate (mappa target e due distrattori) (si veda parte sperimentale). Come visto, l'ipotesi di rappresentazione sovramodale spaziale è stata corroborata dal nostro studio, anche se con le dovute critiche, per cui è necessario svolgere l'esperimento con mappe di complessità similare (si veda discussione dell'esperimento).

In conclusione, nel presente lavoro si è cercato di analizzare e di verificare empiricamente alcune lacune della letteratura e le domande aperte in merito alle finestre temporali, processi di integrazione, sostituzione sensoriale e rappresentazione mentale nelle modalità uditiva e tattile. Come visto, non si sono riscontrate evidenze a favore del fenomeno di integrazione audio-tattile e di sostituzione sensoriale, mentre l'ipotesi della rappresentazione mentale sopramodale è stata confermata. Gli studi svolti, hanno lasciato aperte e hanno aperto delle domande, che richiedono ulteriori studi di approfondimento. Rimane infatti da chiarire il ruolo di mutua influenza e interazione delle modalità tattili e uditiva, e l'effettiva equivalenza funzionale delle due modalità.

Il presente campo di ricerca oltre ad essere rilevante ai fini della comprensione dei singoli fenomeni presi in esame e rispondere quindi, a domande inerenti alla percezione crossmodale in generale, potrebbe avere fini applicativi molto importanti, volti a migliorare la qualità della vita delle persone con deficit uditivi e con principio di sordità. Si consideri, inoltre, che la conferma di una sostituzione sensoriale, potrebbe corroborare l'ipotesi di plasticità audio-tattile anche nella popolazione adulta ed avere applicazioni pratiche che riuscirebbero ad aiutare la qualità della vita di sordi profondi.

5. BIBLIOGRAFIA

Ackroyd, C., Humphrey, N.K., Warrington, E.K. (1974), Lasting effects of early blindness a case study, «*Quarterly Journal of Experimental Psychology*», 26: 114-124

Ahveninen Jyrki , KopcoNorbert , Jääskeläine Iiro P. Psychophysics and neuronal bases of sound localization in humans *Hearing Research* 307 (2014) 86e97

Alain C, Arnott SR, Hevenor S, Graham S, Grady CL. “What” and “where” in the human auditory system”. *Proc Natl Acad Sci* USA. 2001; 98:12301–12306.

Alais D, Burr D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14, 257–262

Alais D. Newell F. N.and Mamassian P. Multisensory Processing in Review: from Physiology to Behaviour. *Seeing and Perceiving* 23 (2010) 3–38.

Alcantara, J. 1., Cowan, R. S. C. Blamey, P. J. & Clark, G. M. 1990a. A comparison of two training strategies for speech recognition with an electrotactile speech processor. *J.speech Hear.Res.*, 33, 195-204.

Alcantara Joseph I . L, Whitford.A. , Blamey P.J. R, Cowan. S.C., and Clark G . M. Speech feature recognition by profoundly hearing impaired children using a multiple-channel electrotactile speech processor and aided residual hearing, J. Acoust. Soc. Am. 88 (3), September 1990b

Alcantara J., Blamey P.J and Clark G. M. (1993). Tactile-Auditory Speech Perception by Unimodally and Bimodally Trained Normal-Hearing Subjects. *J Am Acad Audiol* 4, 98-108

ALCORN, S. Development of the Tadoma method for the deaf-blind. J*ournal of Exceptional Children*,1945, 11, 117-119.

Alcorn S. (1932). The Tadoma method. *Volta Review.* 34, 195-198

Amedi A., Stern W.M., Camprodon J. A. , Bermpoh F., Merabet L., Rotman S., Hemond C., Meijer P. & Pascual-Leone A. Shape conveyed by visualto-auditory sensory substitution activates the lateral occipital complex. *NATURE NEUROSCIENCE* VOLUME 10 [NUMBER 6 [JUNE 2007

Armel KC, Ramachandran VS. Acquired synaesthesia in retinitis pigmentosa. *Neurocase* 1999;5:293–296.

Arnott SR, Binns MA, Grady CL, Alain C. Assessing the auditory dual-pathway model in humans. *Neuroimage*. 2004; 22:401–408.

Backus BT, Haijiang Q. (2007). Competition between newly recruited and preexisting visual cues during the construction of visual appearance. *Vision Research* 47, 919–924.

Banissoni, M., Ponzo, E., and Valvo, A. 1967. Percezione strutturata e trasposizione di forme nelle prime esperienze visive di un cieco dai primi mesi di vita operato in età adulta di osteo-odonto-cheratoprotesi di Strampelli. Annuale di Oftalmologia e Clinica Oculistica 93: 1153–1188.

Banissoni, M., Ponzo, E., and Valvo, A. 1968. Prime esperienze visive di tre cieche dalla nascita operate in età adulta di osteo-odonto-cheratoprotesi di Strampelli: percezione strutturata e trasposizione di forme. Annuale di Oftalmologia e Clinica Oculistica 94: 903–925.

Barlow H. B. . Possible principles underlying the transformation of sensory messages. Sensory Communication, pp. 217–234, 1961

Bastian, H. C. (1880). The brain as an organ of mind. London, England: Kegan & Paul.

Békésy G. v., Human Skin Perception of Traveling Waves Similar to Those on the Cochlea, *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA* VOLUME 27, NUMBER 5, SEPTEMBER, 1955

Békésy G. v., Neural Volleys and the Similarity between Some Sensations Produced by Tones and by Skin Vibrations, *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA* VOLUME 29, NUMBER 10, October, 1957

Békésy G. v. , Sensations on the Skin Similar to Directional Hearing, Beats, and Harmonics of the Ear, *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA* VOLUME 29, NUMBER 4 APRIL. 1957

Békésy G. v. , SIMILARITIES BETWEEN HEARING AND SKIN SENSATIONS, *THE PSYCHOLOGICAL REVIEW*, VOL. 66, No. 1 JANUARY, 1959

Bertelson P, Vroomen J, and Gelder B.VISUAL RECALIBRATION OF AUDITORY SPEECH IDENTIFICATION: A McGurk Aftereffect. *PSYCHOLOGICAL SCIENCE* VOL. 14, NO. 6, NOVEMBER 2003

Blarney, P. J., Alcantara, J. L, Cowan, R. S. C., Galvin, K. L., Sarant, J. Z. & Clark, G. M. 1992. Perception of amplitUde envelope variations of pulsatile electrotactile stimuli. *J Acoust.Soc Am.*. 88, 1765-1772.

Blarney PJ and Clark GM. Psychophysical studies relevant to the design of a digital electrotactile speech processor. *J Acoust SOC Am* 1987;82(1):116-125.

Blauert, J., 1997. Spatial Hearing. The Psychophysics of Human Sound Localization. *MIT Press*, Cambridge, MA.

Blizley Jennifer K., Cohen Yale E. 2013 The what, where and how of auditory-object perception. *Nat Rev Neurosci*, 14 (10):693-707

Bolognini N, Papagno C, Moroni D, Maravita A. Tactile Temporal Processing in the Auditory Cortex *Journal of Cognitive Neuroscience* 22:6, pp. 1201–1211

Bonda, E., Petrides, M., Evans, A., 1996. Neural systems for tactual memories. J. Neurophysiol. 75, 1730 – 173

Boring, E. G. (1942). Sensation and perception in the history of experimental psychology. *New York, NY*: Appleton Century.

Bregman, A. S. Auditory scene analysis: The perceptual organization of sounds. *London: MIT Press*, 1990

Bregman, 1993 Auditory complex scene: hearing in complex environments. DOI:10.1093/acprof:oso/9780198522577.003.0002

Bregman A S, Ahad P A 1996 Demonstrations of auditory scene analysis: The perceptual organization of sound. (Compact disk and booklet). *The MIT Press, Cambridge*, MA

Brockmole, J.R. and Franconeri, S.L. (2009) Introduction. Vis. Cogn. 17,1–9

Brooks P.L. , B.J. Frost, J.L. Mason and K. Chung, Acquisition of a 250-word vocabulary through a tactile vocoder" J. Acoust. Soc. Am., Vol.77, No.4, April 1985

Brooks PL, Frost BJ, Mason JL, and Gibson DM Identification of words and manner features of speech by profoundly deaf subjects using the tactile vocoder . *J Speech Hear Res*

Brooks PL, Frost BJ, Mason JL, Gibson DM : Continuing evaluation of the Queen's University tactile vocoder . 11: Identification of open-set sentences and tracking narrative. *J Rehabil Res Dev* 23(1) :129-138, 1986.

Brooks PL, Frost BJ, Mason JL, Gibson DM (1986). Continuing evaluation of the Queen's University tactile vocoder I:Identification of open set words. *J Rehabil Res Dev* 23(1), 119-128, 1986

Brooks PL, Frost BJ : Evaluation of a tactile vocoder for word recognition . *J Acoust Soc Am* 74(1) :34-39, 1983.

Brooks PL : Comprehension of speech by profoundly deaf and normal-hearing subjects using the Queen's University tactile vocoder . *PhD diss* ., Queen's University, Ontario, Canada, 1984.

Budd TW e Timora JR, Steady state responses to temporally congruent and incongruent auditory and vibrotactile amplitude modulated stimulation, *Int J Psychophysiol.* 2013 Sep;89(3):419-32. doi: 10.1016/j.ijpsycho.2013.06.001

Burk MH, Humen LE, Amos NE, Strauser LE, Effect of Training on Word-Recognition Performance in Noise for Young Normal-Hearing and Older Hearing-Impaired Listeners *EAR & HEARING* / JUNE 2006. Pages 263-278

Cappe C. , Rouiller E.M. , Barone P. Multisensory anatomical pathways. Hearing Research 258 (2009) 28–36

Caetano G and Jousmäki V. (2006). Evidence of vibrotactile input to human auditory cortex. NeuroImage 29(1), 15–28

Calabrese I. and Olivetti Belardinelli M. (1997). Musical Abilities in Deaf Children: Assessment Development and Training. In: A. Gabrielsson (Ed.). *Third Triennal ESCOM Conference: Proceedings*. Uppsala

Calvert, G.A., 2001. Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb. Cortex* 11, 1110 – 1123.

CALVERT GA, BRAMMER MJ, BULLMORE ET, CAMPBELL R, IVERSEN SD, AND DAVID AS. Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*. 10: 2619–2623, 1999.

CALVERT GA, HANSEN PC, IVERSEN SD, AND BRAMMER MJ. Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the bold effect. *Neuroimage* 14: 427–438, 2001.

Calvert GA, Thesen T. Multisensory integration: methodological approaches and emerging principles in the human brain. *J Physiol* Paris.2004 Jan-Jun;98(1-3):191-205.

Carello, C., & Turvey, M. T. (2000). Rotational dynamics and dynamic touch. InM.Heller(Ed.),Touch, representation and blindness (Vol. 13, pp. 27–66). Oxford, England: *Oxford University Press.*

Carney A. E., Osberger M.J., Carney E., Robbins A.M., Renshaw J., and Miyamotc R.T., A comparison of speech discrimination with cochlear implants and tactile aids, *J. Acoust. Soc. Am.* 94(4), October 1993

Caselli, R.J., 1991. Rediscovering tactile agnosia. *Mayo Clin. Proc.* 66, 129 – 142.

Caselli, R.J., 1993. Ventrolateral and dorsomedial somatosensory association cortex damage produces distinct somesthetic syndromes in humans. *Neurology* 43, 762 – 771.7.

Chan JS, Newell FN (2008) Behavioral evidence for task dependent ‘‘what’’ versus ‘‘where’’ processing within and across modalities. *Perception & Psychophysics* 70: 36–49.

Chouvardas V.G., A.N. Miliou, M.K. Hatalis, Tactile displays: Overview and recent advances, *Displays* 29 (2008) 185–194

Connolly Kevin Multisensory perception as an associative learning process. Front Psychol. 2014; 5: 1095.Published online 2014 Sep 26. doi:10.3389/fpsyg.2014.01095

Cowan, R. S. c., Blamey, P. J., Sarant. J. Z., Galvin, K. L. & Clark, G. M. Perception of multiple electrode stimulus patterns: Implications for design of an electrotactile speech processor. *J.Acoust. Soc. Am.*, Vol.89, No.1, January 1991

Cowan, R. S. C ., Alcantara, J. 1.. Blamey. P. J. & Clark. G. M. 1988. Preliminary evaluation of a multichannel electrotactile speech processor. *JAcoust.SocAm*., 83, 2328-2338.

Cowan R.S.C., Blarney J., Galvin K. L.,Sarant J.Z., Alcantara J. I. and Clark G. M. (1990). Perception of sentences, words, and speech features by profoundly hearing-impaired children using a multichannel electrotactile speech processor *J. Acoust. Soc. Am.* 88 (3), 1374-1384

Cowan RSC, Alcantara JI, Whitford LA, Blarney PJ, and Clark GM. Speech perception studies using a multichannel electrotactile speech processor, residual hearing, and lipreading. *J Acoust Soc Am* 1989a;85(6):2593-2607.

Cowan RSC, Blarney PJ, Alcantara JI, Whitford LA, and Clark GM. Speech feature recognition with an electrotactile speech processor. *Aust J Audiol* 1989b;11(2):57-75.

Cox R.M., Johnson J.A., Xu J.; Impact of Advanced Hearing Aid Technology on Speech Understanding for Older Listeners with Mild to Moderate, Adult-Onset, Sensorineural Hearing Loss, *Gerontology* 2014;60:557–568

Cox RM, Alexander GC, Xu J: Development of the Device-Oriented Subjective Outcome (DOSO) Scale. *J Am Acad Audiol*. 2014 September; 25(8): 727–736. doi: 10.3766/jaaa.25.8.3

Cox RM, Alexander GC: The abbreviated profile of hearing aid benefit. *Ear Hear* 1995;16: 176–186

Craig, J. C. (1999). Grating orientation as a measure of tactile spatial acuity. *Somatosensory & Motor Research*, 16, 197-206.

Craik KJM, The Nature of Explanation, *Cambridge University Press*, Cambridge, 1943

Cytowic RE. The man who tasted shapes. Cambridge, *MA: MIT Press*, 2000.

Cytowic RE. Synaesthesia: A Union of the Senses. Cambridge, *MA: MIT Press*, 2002.

Darrow Alice-Ann (1989). Music and the Hearing Impaired: A Review of the Research with Implications for Music Educators. *Applications of Research in Music Education*. 7, 10-12

Darwin, C.J. and Bethell-Fox, C.E. (1977) Pitch continuity and speech source attribution *J.Exp. Psychol. Hum. Percept. Perform.* 3, 665-672

Darwin C J, Carlyon, R P 1995 Auditory grouping. In: B C J Moore, (ed.) *Handbook of perception and cognition: Hearing.* (2nd ed.), Academic Press, London: 387-424

Darwin C J, Carlyon, R P 1997 Auditory grouping *Trends In Cognitive Sciences*, 1, 9

De Filippo Carol Lee and Scott Brian L., A method for training and evaluating the reception of ongoing speech, *J. Acoust. Soc. Am.* 63(4), Apr. 1978

de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, 7(10), 460–467.

Delogu F. , Olivetti Belardinelli M., Palmiero M,Pasqualotto E, Zhao H., Plaisant C., Federici S. : Interactive sonification for blind people exploration of geo-referenced data: comparison between a keyboard-exploration and a haptic-exploration interfaces. Cognitive Processing 7 (Supplement 1) (2006) 178-179.

Delogu F., Palmiero M., Federici S. , Plaisant C., Zhao H., Olivetti Belardinelli M.: Non-visual exploration of geographic maps: Does sonification help?. Disability and Rehabilitation: Assistive Technology 5(3) (2010) 164–174.

Delogu F , Bergmann Tiest WM, Nijboer T C. W., Kappers AM. L.,Postma A Binding in Haptics: Integration of ‘‘What’’ and ‘‘Where’’ Information in Working Memory for Active Touch, PLOS ONE | www.plosone.org 1 February 2013 | Volume 8 | Issue 2 | e55606

Delogu F., Gravina M., Nijboer T., Postma A., Binding “what” and “where” in auditory working memory: An asymmetrical association between sound identity and sound location. *Journal of Cognitive Psychology (*2014) pages 788-798 DOI: 10.1080/20445911.2014.959448

Denham, S.L., Winkler, I., 2006. The role of predictive models in the formation of auditory streams. *J. Physiol.* Paris 100, 154–170.

Deutsch, D. (1982) Grouping mechanism in music. *The Psychology of Music* pp. 99-134 Academic Press

Deutsch D. (1979) Binaural integration of melodic patterns. *Percept. Psychophys.* 25, 399-405

Dijkerman, H. C., & De Haan, E. H. F. (2007). Somatosensory processes subserving perception and action. *Behavioral & Brain Sciences*, 30, 189-201.

Di Lollo, Vincent (2012) The feature-binding problem is an ill-posed problem. *Trends Cogn. Sci.* 16, 317–321

Di Lollo Vincent Response to Holcombe and Clifford: of feature binding and object perception. Volume 16, Issue 8, August 2012, Pages 403 DOI: http://dx.doi.org/10.1016/j.tics.2012.06.012

Di Lollo Vincent Response to Wolfe: feature-binding and object perception Volume 16, Issue 6, June 2012, Pages 308–309

Di Luca M., Ernst M. O. and Backus B. T. (2010). Learning to Use an Invisible Visual Signal for Perception. *Current Biology* 20, 1860–1863 doi:10.1016/j.cub.2010.09.047

Eilers R. E., Widen J. E., and Oller K.D., Assessment techniques to evaluate tactual aids for hearing-impaired subjects, *Journal of Rehabilitation Research and Development* Vol. 25 No . 2 Pages 33—46, Spring 1988

Eimas P. D. and Corbit J. D. (1973). Selective Adaptation of Linguistic Feature Detectors. *Cognitive Psychology* 4, 99-109

Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA. Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron. 2*009; 61:317–329.

Ellis Daniel P. W. Prediction-driven computational auditory scene analysis 1996 *Massachusetts Institute of Technology*

Elsasser, W. M. (1998). Reflections on a theory of organisms. Baltimore, MD: *Johns Hopkins University Press.*

Engel Andreas K. and Singer Wolf Temporal binding and the neural correlates of sensory awareness 2001 *Elsevier Science* PII: S1364-6613(00)01568-0

Engelmann S and Rosov R. J., Tactual Hearing Experiment with Deaf and Hearing Subjects Siegfried, *Research Bulletin* Vol. 14 No. 5, June 1974

Engelmann S and Rosov R. Tactual hearing experiment with deaf and hearing subjects. *Except Child* 1975;41:243-253.

Ernst Marc O. and Bu ¨lthoff Heinrich H.Merging the senses into a robust percept. TRENDS in Cognitive Sciences Vol.8 No.4 April 2004

Feldman Jerome The neural binding problem(s) *Cogn Neurodyn* (2013) 7:1–11 DOI 0.1007/s11571-012-9219-8

Flach, R., & Haggard, P. (2006). The cutaneous rabbit revisited. *Journal of Experimental Psychology: Human Perception & Performance*, 32, 717-732

Fodor J.A. (1983) The Modularity of Mind. An Essay on Faculty Psychology. Cambridge. The *MIT Press.*

Forster, B., Eimer, M., 2004. The attentional selection of spatial and nonspatial attributes in touch, ERP evidence for parallel and independent processes. *Biol. Psychol.* 66, 1–20

Foster JR, Haggard MP: The four alternative auditory feature test (FAAF) – linguistic and psychometric properties of the material with normative data in noise. *Br J Audiol* 1987; 21: 165–174.

Fowler, C.,& Dekle, D. (1991). Listening with eye and hand: Crossmodal contributions to speech perception. *Journal of Experimental Psychology – Human Perception and Performance*, 17, 816–828.

Foxe John J. et al. , Multisensory auditory–somatosensory interactions in early cortical processing revealed by high-density electrical mapping, *Cognitive Brain Research* 10 (2000) 77–83

Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM. (2002). Auditory-Somatosensory Multisensory Processing in Auditory Association Cortex: An fMRI Study. J *Neurophysiol* 88, 540–543, doi:10.1152/jn.00694.2001

Foxe J.J. (2009). Multisensory integration: frequency tuning of audio-tactile integration. *Curr Biol* 19 (9), 373-375 doi: 10.1016/j.cub.2009.03.029

Fujisaki W, Nishida S. A common perceptual temporal limit of binding synchronous inputs across different sensory attributes and modalities. *Proc. R. Soc*. B (2010) 277, 2281–2290 doi:10.1098/rspb.2010.0243

Fujisaki W. e Nishida S. , Audio-tactile superiority over visuo-tactile and audio-visual combinations in the temporal resolution of synchrony perception, *Exp Brain Res* (2009) 198:245–259 DOI 10.1007/s00221-009-1870-x

Galambos, R., Makeig, S., Talmachoff, P.J., 1981. A 40-Hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences of the United States of America* 78 (4), 2643–2647.

Gallace, A., Auvray, M., Tan, H. Z., & Spence, C. (2006). When visual transients impair tactile change detection: A novel case of crossmodal change blindness? *Neuroscience Letters*, 398, 280-285.

Galloni Gloria . Integrazione multisensoriale: la multimodalità e l'insegnamento a distanza. (2003)

Galvin, K. L., Cowan. R. S. C., Sarant, J. Z., Blamey. P. J. & Clark, G. M. 1993. Factors in the development of a training program for use with tactile devices. *Ear Hear*., 14, 118-127.

Galvin, K. L., Mamas, G., Moore. A., Cowan, R. S. C. Blamey. P. J. & Clark, G. M. 1999. A comparison of Tactaid II+ and Tactaid 7 use by adults with a profound hearing impairment. *Ear Hear*. 20,471-482.

Galvin, K. L., Cowan. R. S. C., Blamey. P. J. & Clark, G. M. A Comparison of a New Prototype Tickle Talker™ with the Tactaid 7 *THE AUSTRALIAN AND NEW ZEALAND JOURNAL OF AUDIOLOGY* VOlUME 23 NUMBER 1 MAY 2001 pp 18-36

Galvin, Karyn L.; Cowan, Robert S.C.; Sarant, Julia Z.; Blamey, Peter J.; Clark, Graeme M. Factors in the development of a training program for use with tactile devices *Ear and Hearing*, 14(2), 118-127. (2001)

Gatehouse S, Noble W: The Speech, Spatial and Qualities of Hearing Scale (SSQ). *Int J Audiol* 2004; 43: 85–99.

Gault, R.H., 1926. Touch as a substitute for hearing in the interpretation and control of speech. Arch. *Otolaryngol*. 3, 121 – 135.

Gault, R. H., and Crane, C. W. Tactual patterns from certain vowel qualities instrumentally communicated from a speaker to a subject's fingers. *Journal of General Psychology*, 1928, 1, 353-359.

GAULT, R. H. Hearing through the sense organs of touch and vibration. *Franklin Institute Journal*, 1927, 204, 329-358.

GAULT, R. H. Progress in experiments on tactual interpretation of oral speech. *Journal of Abnormal and Social Psychology*, 1924, 14, 155-159.

Geldard, F. A., & Sherrick, C. E. (1972). The cutaneous "rabbit": A perceptual illusion. *Science*, 178, 178-179.

GESCHEIDER GEORGE A. CUTANEOUS SOUND LOCALIZATION, *Journal of Experimental Psychology* 1965, Vol. 70, No. 6, 617-625

Gescheider George A. , AUDITORY AND CUTANEOUS APPARENT SUCCESSIVENESS, *Journal of Experimental Psychology* 1967, Vol. 73, No. 2, 179-186

Gescheider George A. and Robert K. Niblett, CROSS-MODALITY MASKING FOR TOUCH AND HEARING, *Journal of Experimental Psychology* 1967, Vol. 74, No. 3, 313-320

Gescheider George A. , Ro1e of Phase-Difference Cues in the Cutaneous Analo of Auditory Sound Localization, *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, VOLUME 43, NUMBER 6, 1968

Gescheider GA. (1970). Some comparisons between touch and hearing. IEEE Trans Man-Machine Syst 11, 28–35.

Gescheider GA. (1974). Effects of signal probability on vibrotactile signal recognition. Percept Mot Skills 38, 15–23.

GESCHEIDER GEORGE A. , LAWRENCE C. SAGER and LYDIA J. RUFFOLO, Simultaneous auditory and tactile information processing, *Perception & Psychophysics* 1975, Vol. 18 (3), 209·216

George A. Gescheider, Some Comparisons Between Touch and Hearing, IEEE TRANSACTIONS ON MAN-MACHINE SYSTEMS, MARCH 1970

GIANNOLI GIOVANNI IORIO Il problema di Molyneaux e la "neurologia dei concetti". In Visione, percezione e cognizione nell'età dell'Illuminismo. *Filosofia, estetica, materialismo*, pp. 355-415, Bibliopolis, Napoli, 2005.

Gibson, J. J. (1962). Observations on active touch. *Psychological Review*, 69, 477-491.

Gick B, Jóhannsdóttir KM, Gibraiel D and Mühlbauer J Tactile enhancement of auditory and visual speech perception in untrained perceivers. *J. Acoust. Soc. Am.* 123 4, April 2008

Gick B. & Derrick D. (2009). Aero-tactile integration in speech perception *Nature* 462, 502-504

Gillmeister H. and Eimer M. (2007). Tactile enhancement of auditory detection and perceived loudness. *Brain Res* Jul 30;1160:58-68. Epub 2007 Mar 20

Ginsburg, M.D., Yoshii, F., Vibulsresth, S., Chang, J.Y., Duara, R., Barker, W.W., Boothe, T.E., 1987. Human task-specific somatosensory activation. *Neurology* 37, 1301 – 1308.

Giordano BL, Visell Y, Yao HY, Hayward V, Cooperstock JR, McAdams S. , Identification of walked-upon materials in auditory, kinesthetic, haptic, and audio-haptic conditions, *J Acoust Soc Am*. 2012 May;131(5):4002-12. doi: 10.1121/1.3699205.

GIUDICE N.-C.M. MADISON-P. COSTELLO-J.C. ZHUANG-G.E. LEGGE-X. HU-S. HE, Tactile vision: neuroimaging and brainreorganization in the blind; implications for learning and adaptive-technology, *Proceedings of Technology And Persons With Disabilities Conference*, California State University 2002.

Gobbele´, R., Schu¨ rmann, M., Forss, N., Juottonen, K., Buchner, H., Hari, R., 2003. Activation of the human posterior parietal and temporoparietal cortices during audiotactile interaction. *NeuroImage* 20, 503 – 511.

González Vargas José , Audio Aided Electro-Tactile Perception Training for Finger Posture Biofeedback, 30th *Annual International IEEE EMBS Conference Vancouver*, British Columbia, Canada, August 20-24, 2008

Gray Charles M. The Temporal Correlation Hypothesis of Visual Feature Integration: Still Alive and Well *Neuron*

Gregory, R. (1974), Concepts and mechanisms of perception ,London,

Griffiths, T. D. & Warren, J. D. The planum temporale as a computational hub. Trends *Neurosci*. 25, 348–353 No.7 July 2002

Griffiths Timothy D., Warren Jasonj D., 2004 What is an auditory object? *Nat Rev Neurosci*, 5, 887-892

Grimault, N., Bacon, S.P., Micheyl, C., 2002. Auditory stream segregation on the basis of amplitude-modulation rate. *J. Acoust. Soc. Am*. 111, 1340.

Grunwald, M. (Ed.). (2008). Human haptic perception: Basics and applications. *Basel, Switzerland: Birkhauser.*

Guelke, R. W., and Huyssen, R. M. J.Development of apparatus for the analysis of sound by the sense of touch. *Journal of the Acoustical Society of America*, 1959, 31, 799-809.

Guest Steve , Catmur Caroline , Lloyd Donna and Spence Charles Audiotactile interactions in roughness perception Exp Brain Res (2002) 146:161–171 DOI 10.1007/s00221-002-1164-z

Haijiang Q, Saunders JA, Stone RW, Backus BT (2006). Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. *Proc Natl Acad Sci* U S A 103, 483–488.

Harrison S, Backus B (2010). Disambiguating Necker cube rotation using a location cue: what types of spatial location signal can the visual system learn? *Journal of Vision* 10, 23.

Heller, M. A., Calcaterra, J. A., Burson, L. L., & Green, S. L. (1997). The tactual horizontal–vertical illusion depends on radial motion of the entire arm. *Perception & Psychophysics*, 59, 1297-1311.

Helson, H., & King, S. M. (1931). The tau effect: An example of psychological relativity. *Journal of Experimental Psychology*, 14, 202-217.

Hill, K.T., Bishop, C.W., Yadav, D., Miller, L.M., 2011. Pattern of BOLD signal in auditory cortex relates acoustic response to perceptual streaming. *BMC Neurosci*. 12, 85.

Holcombe, A.O. and Clifford, C.W.G. (2012) Failures to bind spatially coincident features: comment on Di Lollo. *Trends Cogn. Sci. 1*6, 402

Holcombe, AO (2009). The Binding Problem. . In E. Bruce. Goldstein (Ed.), The Sage Encyclopedia of Perception. Sage.

Holger F. Sperdin, Céline Cappe, Micah M. Murray, Auditory-somatosensory multisensory interactions in humans: dissociating detection and spatial discrimination, *Neuropsychologia*, 2010 Nov;48(13):3696-705. doi: 10.1016/j.neuropsychologia.2010.09.001.

Homme Bernhard Event files: feature binding in and across perception and action *TRENDS in Cognitive Sciences*

Hueber T., Benaroya E., Chollet G., Denby B., Dreyfus G.,. Development of a Silent Speech Interface Driven by Ultrasound and Optical Images of the Tongue and Lips. *Speech Communication, Elsevier* : North-Holland, 2010, 52 (4), pp.288.

Inui, K., Wang, X., Tamura, Y., Kaneoke, Y., Kakigi, R., 2004. Serial processing in the human somatosensory system. *Cereb. Cortex* 14, 851–857.

Iwamura, Y., 1998. Hierarchical somatosensory processing. Curr. Opin. Neurobiol. 8, 522–528.

Jacobs L, Karpik A, Bozian D, Gothgen S. Auditory-visual synesthesia: sound-induced photisms. *Arch Neurol* 1981;38: 211–216.

Jain A, Backus BT (2010). Experience affects the use of ego-motion signals during 3D shape perception. *Journal of Vision* 10.

Jain A, Fuller S, Backus BT (2010). Absence of cue-recruitment for extrinsic signals: sounds, spots, and swirling dots fail to influence perceived 3D rotation direction after training. *PLoS One* 5, e13295.

James TW, Kim S, Fisher JS. The neural basis of haptic object processing. *Can J Exp Psychol* 61: 219–229, 2007.

Jensen NS, Akeroyd MA, Noble W, Naylor G: The Speech, Spatial and Qualities of Hearing Scale (SSQ) as a benefit measure: Fourth NCRAR International Conference, Portland/OR, 2009

Johnson, K. O., & Phillips, J. R. (1981). Tactile spatial resolution. I. Two-point discrimination, gap detection, grating resolution, and letter recognition. *Journal of Neurophysiology*, 46, 1177-1191.

Jousmäki V. and Hari R. (1998) Parchment-skin illusion: sound-biased touch. Current Biology, 8 (6), 190.

Kappers, A. M. L. (2003). Large systematic deviations in a bimanual parallelity task: Further analysis of contributing factors. *Acta Psychologica*, 114, 131-145.

Kappers, A. M. L. (2007). Haptic spatial processing: Allocentric and egocentric reference frames. *Canadian Journal of Experimental Psychology*, 61, 208-218.

Kassubba T., Menz M. M. , Ro B. and Siebner H. R. (2013). Multisensory Interactions between Auditory and Haptic Object Recognition. *Cerebral Cortex* 23, 1097-1107.

Killion Mead C. , Niquette,Patricia A. and Gudmundsen Gail I. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners *J Acoust Soc Am* Vol 116, N4, 2004

Kim J-K, Zatorre RJ. 2011. Tactile-auditory shape learning engages the lateral occipital complex. *J Neurosci*. 31:7848--7856

Kim JK e Zatorre RJ., Can you hear shapes you touch?, *Exp Brain Res.*2010 May;202(4):747-54. doi: 10.1007/s00221-010-2178-6.

KIRMAN JACOB H. , TACTILE COMMUNICATION OF SPEECH: A REVIEW AND AN ANALYSIS *Psychological Bulletin* 1973, Vol. 80, No. 1, 54-74

Klatzky L. , Carnegie Mellon University Haptic Processing of the Location of a Known Property: Does Knowing What You've Touched Tell You Where It Is? *Canadian Journal of Experimental Psychology*, 2004, 58:1, 32-45

Klatzky, R. L., & Lederman, S. J. (1993). Toward a computational model of constraint-driven exploration and haptic object identification. *Perception*, 22, 597-621

Klatzky, R. L., & Lederman, S. J. (2003). Representing spatial location and layout from sparse kinesthetic contacts. Journal of Experimental Psychology: *Human Perception & Performance*, 29, 310-325.

Klatzky, R. L., Lederman, S. J., & Metzger, V. A. (1985). Identifying objects by touch: An "expert system." *Perception & Psychophysics*, 37, 299-302

Klatzky RL, Lederman SJ. Identifying objects from a haptic glance. *Percept Psychophys* 57: 1111–1123, 1995.

Klatzky RL, McCloskey B, Doherty S, Pellegrino J, Smith T. Knowledge about hand shaping and knowledge about objects. *J Mot Behav* 19: 187–213, 1987.

Kringlebotn, M.Experiments with some visual and vibrotactile aids for the deaf. *American Annals of the deaf*, 1968, 113, 311-317.

Kubovy M., van Valkenburg D. 2001 Auditory and Visual Objects. *Cognition*, 80, 97-126

Kumar S, Stephan KE, Warren JD, Friston KJ, Griffiths TD. Hierarchical processing of auditory objects in humans. *PLoS Computat Biol.* 2007; 3:e100.

Lagarde J, Zelic G, Mottet D., Segregated audio–tactile events destabilize the bimanual coordination of distinct rhythms, *Exp Brain Res* (2012) 219:409–419 DOI 10.1007/s00221-012-3103-y

Lamme Victor A.F., Roelfsema Pieter R., 2000 The distinct modes of vision offered by feeedforward and recurrent processing. *Trends in Neur* 23, 11, 571-579

Langdon, A.J., Boonstra, T.W., Breakspear, M., 2011. Multi-frequency phase locking in human somatosensory cortex. *Progress in Biophysics and Molecular Biology* 105. (1–2), 58–66. 10: 2619–2623, 1999. 20:361–366

Larsen, E., Iyer, N., Lansing, C.R., Feng, A.S., 2008. On the minimum audible difference in direct-to-reverberant energy ratio. *J. Acoust. Soc.* Am. 124, 450e461.

Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19, 342-368.

Lederman, S. J., & Klatzky, R. L. (2009). Human haptics. In L. R. Squire (Ed. in Chief), *Encyclopedia of neuroscience* (Vol. 5, pp. 11-18). San Diego: Academic Press.

Lederman and KlatzKy, Haptic perception: a tutorial, *Attention, Perception, & Psychophysics* 2009, 71 (7), 1439-1459 doi:10.3758/APP.71.7.1439

Lenay, C., Gapenne, O., Hanneton, S., Marque, C., & Genouëlle, C. (2003). Sensory Substitution : Limits and Perspectives. *Touching for Knowing Cognitive Psychology of Haptic Manual Perception*, *19*, 275–292.

Lessell S, Cohen MM. Phosphenes induced by sound. *Neurology* 1979;29:1524 –1527.

Leva¨nen, S., Hamdorf, D., 2001. Feeling vibrations: enhanced tactile sensitivity in congenitally deaf humans. *Neurosci*. Lett. 301, 75 – 77

Lewis, J. W. et al. (2004) Human brain regions involved in recognizing environmental sounds. *Cereb. Cortex* 14

Locke, J. (1688), Extrait d'un livre anglais qui n'est pas encore publié, intitulé Essai philosophique concernant l'entedement, où l'on montre quelle est l'étendue de nos connaissance certaines, et la manière dot nous y parvenons, «*Bibliotèque universelle et Historique*», 8: 49-14

Loizou Philipos C. Mimicking the Human Ear. An overview of Signal-Processing Strategies for Converting Sound into Electrical Signals in Cochlear Implants. *IEE SIGNAL PROCESSING MAGAZINE* September 1998

Lu¨tkenho¨ ner, B., Lammertmann, C., Simoes, C., Hari, R., 2002. Magnetoencephalographic correlates of audiotactile interaction. *NeuroImage* 15, 509 – 522

Lunghi C., Morrone M C. and Alais D. (2014) Auditory and Tactile Signals Combine to Influence Vision during Binocular Rivalry. *Journal of Neuroscience* 34 (3) 784-792; doi: https://doi.org/10.1523/JNEUROSCI.2732-13.2014

Luria AR. The mind of a mnemonist. *New York: Basic Books*, 1968.

Lynch M.P., MS, Eilers R.E., Oller D.K. and Laoie L., (1988) Speech perception by congenitally deaf subjects using an electrocutaneous vocoder *Journal of Rehabilitation Research and Development* Vol . 25 No. 3 Pages 41-50

Lynette A. Jones, Nadine B. Sarter Tactile Displays: Guidance for Their Design and Application *Human Factor* Volume: 50 issue: 1, page(s): 90-111

Lynn C. Robertson Binding, spatial attention and perceptual awareness *Nature Reviews Neuroscience* 4, 93-102 (February 2003) | doi:10.1038/nrn1030

Macaluso E, Driver J. Multisensory spatial interactions: a window onto functional integration in the human brain.*Trends Neurosci.*2005 May;28(5):264-71.

MACALUSO E, FRITH CD, AND DRIVER J. Modulation of human visual cortex by crossmodal spatial attention. *Science* 289: 1206–1208, 2000.

Macpherson, E.A., Sabin, A.T., 2007. Binaural weighting of monaural spectral cues for sound localization. *J. Acoust. Soc. Am.* 121, 3677e3688.

Maeda F., Kanai R., Shimojo S. (2004). Changing pitch induced visual motion illusion. *Current Biology* 14(23), 990.

Maeder PP, et al. Distinct pathways involved in sound recognition and localization: a human fMRI study. Neuroimage. 2001; 14:802–816.

Massaro, D. (1987). Speech perception by ear and eye: A paradigm for psychological inquiry. *Hillsdale, NJ: Erlbaum.*

Massaro, D.W. & Cohen M.M. (1983) Evaluation and integration of visual and auditory information in speech perception *J Exp Psychol Human* 9 (5), 753-771.

Massaro D.W., Cohen M.N., Schwartz W., Vanderhyden S., Meyer H. Facilitating Speech Understanding for Hearing-Challenged Perceivers in Face-to- Face Conversation and Spoken Presentations, Paper presented to *the1st International Conference on Technology for Helping People with Special Needs (ICTHP-2013).*

McGurk H. and MacDonald J. (1976). Hearing lips and seeing voices *Nature*, 264(5588), 746–748.

Micheyl C, Kreft H, Shamma S, Oxenham AJ. Temporal coherence versus harmonicity in auditory stream formation. *J Acoust Soc Am*. 2013; 133:EL188–EL194

Middlebrooks, J.C., Green, D.M., 1991. *Sound localization by human listeners. Annu. Rev. Psychol*. 42, 135e159

Miller G. A. and Nicely P.E. (1955). An Analysis of Perceptual Confusions Among Some English Consonants. *J. Acoust. Soc. Am.* 27, 338. doi: http://dx.doi.org/10.1121/1.1907526

Ming V.L, and Holt LL.Efficient coding in human auditory perception *J. Acoust. Soc. Am*. 126 3, 2009

Mishkin, M., 1979. Analogous neural models for tactual and visual learning. *Neuropsychologia* 17, 139–151

Molyneux, W. (1692), Dioptrica Nova: A treatise of Dioptricks, in Two Parts, Wherein the Various Effects and Appearances of Spherick Glasses, Both Convex and concave, Single and Combined, in Telescopes and Microscopes, Togheter with their Usefulness in many Concerns of Humane Life, are Explained, London, Benj. Tooke.

Moore, B.C.J., 1997. An Introduction to the Psychology of Hearing, fourth ed. Academic Press, San Diego, CA.

Müller J. (1838) Elements of Physiology: https://archive.org/details/Muller1838na12J-b

Navarra J, Soto-Faraco S. and Spence C., Adaptation to audiotactile asynchrony, Neuroscience Letters, Volume 413, Issue 1, 8 February 2007, Pages 72–76

Ohara S., Wang L. , Ku Y., Lenz FA, Hsiao SS , Hong B, Zhou YD , Neural activities of tactile cross-modal working memory in humans: an event-related potential study, Neuroscience 152 (2008) 692–702

Oiler DK, Payne SI,, Gavin WJ : Tactual speech perception by minimally-trained deaf subjects . *J Speech Hear Res* 23(4) :769-778, 1980.

Olivetti Belardinelli: Accessibility of Sonificated Spatial Information to Blind People, *FORUM ACUSTICUM* 2011, 27. June - 1. July, Aalborg, Proceedings

Olivetti Belardinelli M., Delogu F., Palmiero M., Federici S., Mele ML., Fiacchi S., Rizza A.: Audio-tactile

exploration and mental representation of sonificated maps by blind subjects: A review and experimentation according to objective, subjective and interactive perspectives. *Universal Access in the Information Society* (submitted)

Olivetti Belardinelli M., Delogu F. , Palmiero M., Federici S., Pasqualotto E., Zhao H. Plaisant: Interactive sonification of geographical maps: a behavioural study with blind subjects. Proc. 2007 *15th Conference of the European Society for Cognitive Psychology*.

Olivetti Belardinelli M. 2011 Looking for Effective Accessibility of Sonificated Spatial Information to Blind People: Four Experiments. *FORUM ACUSTICUM*, Aalborg

Owens, E ., Kessler, D . K., Raggio, M . W., and Schubert E, . O. (1985). "Analysis andrevision of the minimal auditory capabilities battery", *Ear Hear.* 6(6), 280-287.

O'Regan, J. K., & Noë, a. (2001). A sensorimotor account of vision and visual consciousness. *The Behavioral and Brain Sciences*, 24(5), 939–973; discussion 973–1031. doi:10.1017/S0140525X01000115

O'Regan (2011) Why Red Doesn't Sound Like a Bell: Understanding the feel of consciousness *Oxford University Press*

O'REGAN J. KEVIN AND NOË ALVA, WHAT IT IS LIKE TO SEE: A SENSORIMOTOR THEORY OF PERCEPTUAL EXPERIENCE *Synthese* 129: 79–103, 2001

Pasquinelli Elena Conflitti Intersensoriali. Luglio 2003

Pasquinelli Elena LA COERENZA DELLA PERCEZIONE E LA COSTRUZIONE DI UNITÀ MULTISENSORIALI 2003

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. The processing of temporal pitch and melody information in auditory cortex. Neuron. 2002; 36:767–776

Peterson, G. E. & Lehiste. 1. 1962. Revised CNC lists for auditory tests. *J.Speech Hear.Dis*., 27, 62-70.

Pickett, J. M., and Pickett, B. H. Communication of speech sounds by tactual vocoder . *Journal of Speech and Hearing Research* 1963, 6, 207-222.

Pickett, J. M., Speech Communication for the Deaf: Visual, Tactile, and Cochlear-Implant. *Journal of Rehabilitation Research and Development*, Vol. 23 No. 1 1986. pages 95-99

Plant, G. 1998. Training in the use of a tactile supplement to lipreading: A long-tenn case study. Ear Hear., 19, 394-406.

Pressnitzer D, Sayles M, Micheyl C, Winter IM. Perceptual organization of sound begins in the auditory periphery. *Curr Biol.* 2008; 18:1124–1128.

Proulx Michael J. , Brown David J. , Pasqualotto Achille , Meijer Peter Multisensory perceptual learning and sensory substitution *Neuroscience and Biobehavioral Reviews* xxx (2012) xxx–xxx

Purdy KA, Lederman SJ, Klatzky RL (2004) Haptic processing of the location of a known property: Does knowing what you've touched tell you where it is? *Canadian Journal of Psychology* 58: 32–45.

Rauschecker JP, Tian B. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc *Natl Acad Sci USA*. 2000; 97:11800–11806

Reed, C. L., Klatzky, R. L., & Halgren, E. (2005). What vs. where in touch: An fMRI study. *Neuroimage*, 25, 718–726. doi:10.1016/j.neuroimage.2004.11.044

Reed, C. M., Durlach, N. I., Delhorne, L. A., Rabinowitz, W. M. & Grant, K. W. 1989. Research on tactual communication of speech: Ideas, issues, and findings. *Volta Review*. 91(5),65-78.

Reed, C. M. & Delhorne, L. A. 1995. Current results of a field study of adult users of tactile aids. Seminars in Hearing, 1995

Reed, E. Courtenay Wilson, Tactile Communication of Speech, Chapter 22. *Tactile Communication of Speech*

Reed C.M., Durlach N. I., Braida L.D. Research on tactile communication of speech: a review, *American Speech-Language-Hearing Association*, 1982 MD 20852

Reed C.M. , W. M. Rabinowitz, N. I. Durlach, and L. D. Braida, S. Conway-Fithian and M. C. Schultz Research on the Tadoma method of speech communication, *J. Acoust. Soc. Am.* 77 (1), January 1985

Rees, A., Green, G.G.R., Kay, R.H., 1986. Steady-state evoked-responses to sinusoidally amplitude-modulated sounds recorded in man. *Hearing Research* 23 (2), 123–133.

Renier LA, Anurova I, De Volder AG, Carlson S, VanMeter J, Rauschecker JP. 2009. Multisensory integration of sounds and vibrotactile stimuli in processing streams for ‘‘what’’ and ‘‘where’’. *J Neurosci.* 29:10950—10960

Repp Bruno H. THE ROLE OF PSYCHOPHYSICS IN UNDERSTANDING SPEECH PERCEPTION, *HASKINS LABORATORIES: Status Report on Speech Research* SR-86/87 (1986)

Revesz, G. (1950). Psychology and art of the blind. London, England*: Longmans, Green*.

Richardson BL, Symmons MA Vibrotactile devices for the deaf: are they out of touch? *Clark and Cowan, International Cochlear Implant, Speech and Hearing Symposium*

Riesenhuber, M., Poggio, T. 1999. The role of neural mechanisms of attention in solving the binding problem. *Neuron 24* (September), 19-29

Rinott Michael Audio-Tactile interactions with mobile devices *Interaction Design Institute Ivrea*, May 2004

Ro T. , Hsu J., Yasar N.E., Elmore L.C., Beauchamp M.S. (2009). Sound enhances touch perception. *Exp Brain Res* 195, 135–143. DOI 10.1007/s00221-009-1759-8

Röder B, Pagel B, Heed T. The implicit use of spatial information develops later for crossmodal than for intramodal temporal processing. *Cognition.*2013 Feb;126(2):301-6. doi: 10.1016/j.cognition.2012.09.009.

Roland, P.E., O’Sullivan, B., Kawashima, R., 1998. Shape and roughness activate different somatosensory areas in the human brain. *Proc. Natl. Acad. Sci. U. S. A*. 95, 3295 – 3300.

Ronnberg J, Andersson U, Lyxell B, Spens K., Vibrotactile Speech Tracking Support: Cognitive Prerequisites *Journal of Deaf Studies and Deaf Education* 3:2 Spring 1998

Rosenblatt F: Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Washington, D.C.*:SpartanBooks*;1961

Roskie Adina L.s The Binding Problem *Neuron*, Vol. 24, 7–9, September, 1999

Sacks, O. (1995), An Antropologist on Mars: seven paradoxical tales, New York, *Alfred A. Knopf*

Sanabria D, Soto-Faraco S and Spence C, Spatiotemporal interactions between audition and touch depend on hand posture, *Exp Brain Res* (2005) 165: 505–514 DOI 10.1007/s00221-005-2327-5

Sarant J, Cowan R., Blamey P., Galvin K and Clark G., Within-subject Comparison of Speech Perception Benefits for Congenitally Deaf Adolescents with an Electrotactile Speech Processor and a Cochlear Implant, *J Am Acad Audiol* 7 : 63-70 (1996)

Sato M., Cavé C., Ménard L. and Brasseur A. (2010). Auditory-tactile speech perception in congenitally blind and sighted adults. *Neuropsychologia* 48, 3683–3686.

Schmidt T Perception: The Binding Problem and the Coherence of Perception 2009 *Elsevier*

Schürmann, Martin & Caetano, Gina & Jousmäki, Veikko & Hari, Riitta. (2004). Hands help hearing: Facilitatory audiotactile interaction at low sound-intensity levels. *J.Acoust. Soc. Am.* 115(2), 830-832. ISSN 1520-8524.

Schwartz J, Grimault N, Hupe JN, Moore B C. J. and Pressnitzer D. Multistability in perception: binding sensory modalities, an overview *Phil. Trans. R. Soc. B* (2012) 367, 896–905 doi:10.1098/rstb.2011.0254

Scilley P, . L. (1980). "Evaluation of an auditory prosthetic device for the profoundly deaf," Master's thesis, Queen's University, Kingston, Ontario, *Canada Seminars in Hearing.* 16,305-315.

Sekuler R, Sekuler AB, Lau R (1997). Sound alters visual motion perception [letter]. *Nature* 385, 308.

Senkowski D, Schneider TR, Foxe JJ and Engel AK Crossmodal binding through neural coherence: implication for multisensory processing. *Trends in Neurosciences* Vol.31 No.8 2008 Elsevier Ltd.doi:10.1016/j.tins.2008.05.002

Shadlen Michael N. and Movshon J. Anthony Synchrony Unbound: A Critical Evaluation of the Temporal Binding Hypothesis *Neuron*, Vol. 24, 67–77, September, 1999, Cell Press

Shamma SA, Elhilali M, Micheyl C.Temporal coherence and attention in auditory scene analysis.*Trends Neurosci.* 2011 Mar;34(3):114-23. doi: 10.1016/j.tins.2010.11.002. Epub 2010 Dec 31. 20:361–366

Shamma SA, Micheyl C. Behind the scenes of auditory perception. *Curr Opin Neurobiol.* 2010

Shams, L., Kamitani, Y., and Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature* 408, 788.

Sharon Zmigrod, Leor Zmigrod Zapping the gap: Reducing the multisensory temporal binding window by means of transcranial direct current stimulation (tDCS) *Consciousness and Cognition* 35 (2015) 143–149 –65 (discussion 65-70).

Sharon Zmigrod, Michiel Spapé, Bernhard Hommel Intermodal event files: integrating features across vision, audition, taction, and action, *Psychological Research* (2009) 73:674–684 DOI 10.1007/s00426-008-0163-5

Shaw, E.A., 1966. Earcanal pressure generated by a free sound field. *J. Acoust. Soc. Am.* 39, 465e470.

Sherrick, C. E., & Rogers, R. (1966). Apparent haptic movement. *Perception & Psychophysics*, 1, 175-180

Sherrick, C.E., 1984. Basic and applied research on tactile aids for deaf people: progress and prospects. *J. Acoust. Soc. Am.* 75, 1325 – 1342.

Sherrington, C. S. (1906). On the proprioceptive system especially in its reflex aspect. *Brain*, 29, 467–482.

Shimono, K., Higashiyama, A., & Tam, W. J. (2001). Location of the egocenter in kinesthetic space. *Journal of Experimental Psychology: Human Perception & Performance*, 27, 848-861.

Shull Peter B and Damian Dana D. Haptic wearables as sensory replacement, sensory augmentation and trainer – a review, *Shull and Damian Journal of NeurEngineering and Rehabilitation* (2015) 12:59 DOI 10.1186/s12984-015-0055-z

Sicilley, P. L. (1980) Evaluation of a vibrotactile auditory prosthetic device for the profoundly deaf. [unpublished *Master's thesis*] Queen's University, Kingston, Canada

Simner J, Holenstein E. Ordinal linguistic personification as a variant of synesthesia. *J Cogn Neurosci* 2007;19:694 –703.

Simner J, Mulvenna C, Sagiv N, et al. Synaesthesia: the prevalence of atypical cross-modal experiences. *Perception* 2006;35: 1024–1033.

Singer Wolf Neuronal Synchrony: A Versatile Code Review for the Definition of Relations? *Neuron*, Vol. 24, 49–65, September, 1999, Copyright 1999 by Cell Press

Singh, P.G., 1987. Perceptual organization of complex-tone sequences: a tradeoff between pitch and timbre? J. Acoust. Soc. Am. 82, 886–889.

Smallman, H.S., Fine, I., MacLeod, D. (2000), Pre- and post-operative characterization of visual function after the removal of bilateral congenital cataracts in adulthood, «*Society for Neuroscience Abstracts*», 17: 825

Snyder, A.Z., 1992. Steady-state vibration evoked-potentials—description of technique and characterization of responses. *Electroencephalography and Clinical Neurophysiology* 84 (3), 257–268.

Snyder Joel S., Gregg Melissa K., Weintraub David M., Alain C., 2012 Attention, awareness, and the perception of auditory scenes *Frontiers in Psychology* 3, 15, 1-17

Soto-Faraco S, Spence C. and Kingstone A. (2004). Congruency effects between auditory and tactile motion: Extending the phenomenon of cross-modal dynamic capture. *Cogn Affect Behav Ne* 4 (2), 208-217.

Soto-Faraco S. and Deco G. (2009).Multisensory contributions to the perception of vibrotactile events, *Behav Brain Res.* 2009 Jan 23;196(2):145-54. doi: 10.1016/j.bbr.2008.09.018. Epub 2008 Sep 30

Spence, C., & Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Current Biology*, 13(13), R519–R521.

Spence Charles and Zampini Massimiliano, Auditory Contributions to Multisensory Product Perception, *ACTA ACUSTICA UNITED WITH ACUSTICA* Vol. 92 (2006) 1009–1025

Spence Charles Crossmodal correspondences: A tutorial review. *Atten Percept Psychophys* (2011) 73:971–995 DOI 10.3758/s13414-010-0073-7

Spence Charles Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Ann. N.Y. Acad. Sci.* 1296 (2013) 31–49 C

Spence C, Deroy O. How automatic are crossmodal correspondences? *Conscious Cogn.* 2013 Mar;22(1):245-60. doi: 10.1016/j.concog.2012.12.006.

Spence KE and Plant G, A TACTUAL "HEARING" AID FOR THE DEAF, *STL-QPSR*, Vol. 24 (1), 1983. Pages 052-056.

Stacey P. C. and Summerfield A. Q., Comparison of Word-, Sentence-, and Phoneme-Based Training Strategies in Improving the Perception of Spectrally Distorted Speech, *Journal of Speech, Language, and Hearing Research* Vol. 51, 526–538, April 2008

Stein,B.E.,& Meredith,M.A.(1990).Multisensory integration. Neural and behavioral solutions for dealing with stimuli from different sensory modalities. *Annals of the New York Academy of Sciences*, 608, 51

Stein,B.E.,& Meredith,M.A.(1993). The Merging of the Senses. Cambridge, *MA:MIT Press*

Stein,B.E.(1998).Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Experimental Brain Research*, 123(1–2), 124–135.

STEIN BE, LAURIENTI PJ, WALLACE MT, and. STANFORD TR, Multisensory Integration *Encyclopedia of the Human Brain.* Copyright 2002, Elsevier Science (USA). Volume 3

Stiles, N. R. B., & Shimojo, S. (2015). Sensory Substitution. *International Journal of Neuroscience*, 19(1-4), 1–15.doi:10.3109/00207458309148639

Stoet , G. & Hommel, B. (2002) Interaction between feature binding in perception and action In Common Mechanisms in Perception and Action: Attention and Performance XIX (Prinz, W. and Hommel, B. eds), pp. 538–552, *Oxford University Press*

Stoet, G. and Hommel, B. (1999) Action planning and the temporal binding of response codes. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 1625–1640

Strutt, J.W., 1907. On our perception of sound direction. *Philos. Mag.* 13, 214e232.

Suzuki Y., Gyoba J., Sakamoto S., Selective effects of auditory stimuli on tactile roughness perception, *Brain Res*., 2008 Nov 25;1242:87-94. doi: 10.1016/j.brainres.2008.06.104. Epub 2008 Jul 2.

Tajadura-Jiménez A, Kitagawa N, Väljamäe A, Zampini M, Murray MM, Spence C. , Auditory-somatosensory multisensory interactions are spatially modulated by stimulated body surface and acoustic spectra. *Neuropsychologia* 2009 Jan;47(1):195-203. doi: 10.1016/j.neuropsychologia.2008.07.025. Epub 2008 Aug 6.

Tan H., Durlach N. L., Reed C. M. , and Rabinowitz W. M., Information transmission with a multifinger tactual display, *Perception & Psychophysics* 1999,61 (6), 993-1008

Tan H. and Rabinowitz W.M. A new multi-finger tactual display Proceedings of the Dynamic Systems and Control Division, *DSC* Vol, 58; Book No. G01023- 1996

Thome R, Offutt S, Bagley L, Weaver A, Page J, Tyler MT Tactile Auditory Sensory Substitution. BME 200/300 *University of Wisconsin* - Madison, December 13, 2006

Till R. Schneider, Simone Lorenz, Daniel Senkowski, and Andreas K. Enge Gamma-Band Activity as a Signature for Cross-Modal Priming of Auditory Object Recognition by Active Haptic Exploration *The Journal of Neuroscience*, February 16, 2011, 31(7):2502–2510

Tobimatsu, S., Zhang, Y.M., Kato, M., 1999. Steady-state vibration somatosensory evoked potentials: physiological characteristics and tuning function. *Clinical Neurophysiology* 110 (11), 1953–1958.

Tolman, Edward C. (July 1948). "Cognitive maps in rats and men". *Psychological Review*. 55 (4): 189–208. PMID 18870876. doi:10.1037/h0061626

Tonnison, B . (1974). "National acoustic laboratories: Standardization of CID everyday sentence tests", *N . A. L. Publication*, Sydney, Australia

Treisman, A. (1996) The binding problem. *Curr. Opin. Neurobiol.* 6, 171–178

Treisman, A. & Gelade, G. (1980) A feature-integration theory of attention *Cognitive Psychology*, 12,97–136

Treisman, A. & Schmidt, H. (1982). Illusory conjunctions in the perception of objects., 14, 107-141

Treisman, A. Solutions to the Binding Problem: Review Progress through Controversy Summary and Convergence:Review. *Neuron*, Vol. 24, 105–110, September, 1999

Turvey, M. T., & Carello, C. (2011). Obtaining information by dynamic (effortful) touching. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 366, 3123–3132. doi:10.1098/rstb.2011.0159

Turvey,M. T.,& Shaw, R.E. (1995).Toward an ecological physics and a physical psychology. In R.Solso & D.Massaro(Eds.),*The science of the mind: 2001 and beyond* (pp. 144–169). Oxford, England: Oxford University Press.

Turvey Michael T. & Fonseca Sérgio T. (2014) The Medium of Haptic Perception: A Tensegrity Hypothesis, *Journal of Motor Behavior*, 46:3, 143-187, DOI:10.1080/00222895.2013.798252

Umezu, H., Torii, S., Uemura, Y. (1975), Post operative formation of visual perception in the early blind. «*An International Journal of Psychology in the Orient*», 18: 171-186

Ursino M, Cuppini C, Magosso E Neurocomputational approaches to modelling multisensory integration in the brain: a review. *Neural Netw*. 2014 Dec;60:141-65. doi:10.1016/j.neunet.2014.08.003.

Van Boven, R. W., Ingeholm, J. E., Beauchamp, M. S., Bikle, P. C., & Ungerleider, L. G. (2005). Tactile form and location processing in the human brain. *Proceedings of the National Academy of Sciences*, 102, 12601–12605. doi:10.1073/pnas.0505907102

Van der Stoep N, Nijboer TC, Van der Stigchel S, Spence C Multisensory interactions in the depth plane in front and rear space: a review. *Neuropsychologia.*2015 Apr;70:335-49. doi: 10.1016/j.neuropsychologia.2014.12.007.

van Noorden, L., 1975. Temporal Coherence in the Perception of Tone Sequences. *Technical University, Eindhoven* pp. 1–127.

Velik Rosemarie From single neuron-firing to consciousness—Towards the true solution of the binding problem *Neuroscience & Biobehavioral Reviews* 34 (7), 993-1001

Visell Yon Tactile sensory substitution: Models for enaction in HCI, *Interacting with Computers* xxx (2008)

Vliegen, J., Oxenham, A.J., 1999. Sequential stream segregation in the absence of spectral cues. *J. Acoust. Soc. Am.* 105, 339–346.

C Von Der Malsburg (1981) The correlation theory of brain function.*In Models of Neural Networks II: Temporal Aspects of Coding and Information Processing in Biological Systems* (Domany, J.L. et al., eds), pp. 95–119, Springer-Verlag, (reprinted in 1994)

Von Der Malsburg C. (1995) Binding in models of perception and brain function *Current opinion in neurobiology* 5 (4), 520-526

Von der Malsburg C. (1999) The what and why of binding: the modeler's perspective *Neuron* 24 (1), 95-104

Von Schiller P., Die Rauhigkeit als intermodale Erscheinung. *Z Psychol Bd*,1932, 127:265-289.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics*, 72(4), 871–884

Vroomen Jean , van Linden Sabine , Keetels Mirjam, de Gelder Béatrice, Bertelson Paul Selective adaptation and recalibration of auditory speech by lipread information: dissipation *Speech Communication* Volume 44, Issues 1–4, October 2004, Pages 55-61

Vroomen Jean, van Linden Sabine , de Gelder Beatrice , Bertelson Paul Visual recalibration and selective adaptation in auditory–visual speech perception: Contrasting build-up courses. *Neuropsychologia* 45 (2007) 572–577

Wallace Mark T. and Stevenson Ryan A. The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia* 64 (2014) 105–123106

Walker KM, Bizley JK, King AJ, Schnupp JW. Multiplexed and robust representations of sound features in auditory cortex. *J Neurosci*. 2011; 31:14565–14576

Wankhede S. Nilashree Designing visual Speech Training Aids for Hearing Impaired Children *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* Vol. 3, Issue 4, April 2014

Warren JD, Griffiths TD. Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *J Neurosci.* 2003; 23:5799–5804.

Watson C, Dobie RA, Durlacj N, Humes LE, Levitt H, Miller JD, Sherrich CE, Simmons FB, Studebaker GA, Tyler R, Widin GP, Speech-perception aids for hearing-impaired people: Current status and needed research. *J. Acoust. Soc Am* , 90(2). Pt.1, August 1991

Weinstein, S. (1968). Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality. In D. R. Kenshalo (Ed.), *The skin senses* (pp. 195-222). Springfield, IL: Thomas.

Weisenberger JM. Evaluations of single-channel and multichannel tactile aids for the hearing impaired. *J Acoust Soc Am* 1987

Weisenberger, J.M., Miller, J.D., 1987. The role of tactile aids in providing information about acoustic stimuli. *J. Acoust. Soc. Am.* 82, 906 – 916

Weisenberger J.M. and Broadstone S.M. Frank A. Saunders, (1989) Evaluation of two multichannel tactile aids for the hearing impaired, J. Acoust. Soc. Am. 86 (5)

Weisenberger J.M. and Broadstone S.M., Kozma-Spytek Linda Relative performance of single-channel and multichannel tactile aids for speech perception *Journal of Rehabilitation Research and Development* Vol. 28 No. 2, 1991 Pages 45-56

Wessel, D.L. (1979) Timbre space as a musical control structure. *Comp.Mus.* J. 3, 45-52.

Whitney David Neuroscience: Toward Unbinding the Binding Problem. *Curr Biol.* 2009 Mar 24; 19(6): R251–R253.doi:10.1016/j.cub.2009.01.047

Wiener, N., Wiesner, F. E. David, Jr., and Levine, L. Operation "Felix." *Quarterly Progress Report, Research Laboratory of Electronics, Massachusetts Institute of Technology*, Cambridge, 1949-1951.

Wightman, F.L., Kistler, D.J., 1997. Monaural sound localization revisited. *J. Acoust.Soc. Am*. 101, 1050e1063.

Winkler I, Denham SL, Nelken I. Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends Cogn Sci.* 2009; 13:532–540

Wolfe Jeremy M. and Cave Kyle R. The Psychophysical Evidence for a Binding Problem in Human Vision *Neuron*, Vol. 24, 11–17, September, 1999 Cell Press

Wolfe Jeremy M. The binding problem lives on: comment on Di Lollo Volume 16, Issue 6, June 2012, Pages 307–308

Wolf Singer Neuronal Synchrony: A Versatile Code Review for the Definition of Relations? *Neuron*, Vol. 24, 49–65, September, 1999, Copyright 1999 by Cell Press

Wong, T. S. (1977). Dynamic properties of radial and tangential movements as determinants of the haptic horizontal–vertical illusion with an “L” figure. Journal of Experimental Psychology: *Human Perception & Performance*, 3, 151-164.

Yarrow K., Haggard P. and Rothwell J.C. (2008). Vibrotactile–Auditory Interactions are Post-Perceptual. *Perception* 37, 1114-130.

Yau J. M., Olenczak J. B. , Dammann J. F. and Bensmaia S. J. (2009). Temporal frequency channels are linked across audition and touch. *Curr Biol* 19(7), 561–566. doi:10.1016/j.cub.2009.02.013

Yau J. M. , Weber A. I. and Bensmaia S. J. (2010). Separate mechanisms for audio-tactile pitch and loudness interactions. *Front. Psychol., Perception Science*, 1, Article 160 doi: 10.3389/fpsyg.2010.00160

Yau Jeffrey M. , Alison I. Weber and Sliman J. Bensmaia, Separate mechanisms for audio-tactile pitch and loudness interactions, *Frontiers in Psychology | Perception Science* October 2010 | Volume 1 | Article 160

Yau Jeffrey M., Jonathon B. Olenczak, John F. Dammann, and Sliman J. Bensmaia, Temporal frequency channels are linked across audition and touch, *Curr Biol* . 2009 April 14; 19(7): 561–566. doi:10.1016/j.cub.2009.02.013

YAZDI KARIMI A.R., SAZGAR A.A., NADIMI TEHRANI A., FARAMARZI A.A.H., NASSAJ F.E., YAHYAVI SH, Application and usage of tactile aid in Iran. *ARCHIVES OF IRANIAN MEDICINE* 2006 , Volume 9 , Number 4; Pages 344-347

Yost, W.A., Gourevitch, G., 1987. Directional Hearing. *Springer-Verlag*, New York

Zampini Massimiliano et al., Audiotactile temporal order judgments, *Acta Psychologica* 118 (2005) 277–291

Zampini M., Torresan D., Spence C. and Murray MM. Auditory-somatosensory multisensory interactions in front and rear space, *Neuropsychologia* 2007 Apr 9;45(8):1869-77. Epub 2007 Feb 7.

Zatorre, R. J., Bouffard, M. & Belin, P. (2004).Sensitivity to auditory object features in human temporal neocortex. *J. Neurosci.* 24, 3637–3642

Zmigrodv S , Spapé M, Hommel B Intermodal event files: integrating features across vision, audition, taction, and action, *Psychological Research* (2009) 73:674–684 DOI 10.1007/s00426-008-0163-5

6. APPENDICE

Table 3. Summary of information provided by the Tickle Talker.

Speech Feature	Electrical Parameter	Percept	Information Perceived
F_0	Biphasic pulse rate	Roughness of stimulation	Syllable stress
F_2	Electrode stimulation	Stimulation on electrode in range 2–7	Vowel formant
High-frequency information	Stimulation on electrode 8	Stimulation on electrode 8	High-frequency fricative
Amplitude	Biphasic pulse width	Strength of stimulation	Syllable number and stress Vowel duration Consonant manner
Voicing	Stimulation on electrode 1	Stimulation on electrode 1	Voiced or unvoiced

Fig.1 Esempio della tipologia delle informazioni veicolate dai dispositivi. Nella fattispecie, si parla del Tickle Talker del gruppo seguito da Clark (Galvin et al., 1993).

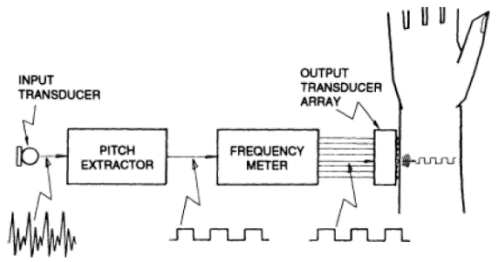


Fig. 2 Esempio di estrazione delle frequenze e trasduzione del codice per la stimolazione tattile (Wankhede, 2014).

TABLE II. Confusion matrix for $S/N = -12$ db and frequency response 200–6500 cps.

	p	t	k	f	θ	s	ʃ	b	d	g	v	ð	z	ʒ	m	n
p	51	53	65	22	19	6	11	2		2	3	3	1	5	8	5
t	64	57	74	20	24	22	14	2	3	1	1	2	1	1	5	1
k	50	42	62	22	18	16	11	4	1	1	1	2			4	2
f	31	22	28	85	34	15	11	3	5		8	8	3		3	
θ	26	22	25	63	45	27	12	6	9	3	11	9	3	2	7	2
s	16	15	16	33	24	53	48	3	5	6	3	1	6	2		1
ʃ	23	32	20	14	27	25	115	1	4	5	3		6	3	4	2
b	4	2	2	18	7	7	1	60	18	18	44	25	14	6	20	10
d	3		1	4	7	4	11	18	48	35	16	24	26	14	9	12
g	3	1	1	1	4	5	7	20	38	29	16	29	29	38	10	9
v		1	1	12	5	4	5	37	20	23	71	16	14	4	14	9
ð		1	4	17	2	3	2	53	31	25	50	33	23	5	13	6
z	6	1	2	2	6	14	8	23	29	27	24	19	40	26	3	6
ʒ	3	2	2	1		6	7	7	30	23	9	7	39	77	5	14
m		1			1	1		11	3	6	8	11		1	100	60
n	1			1		1		2	2	6	7	1	1	9	84	145

Fig. 3 Matrice di confusione della percezione delle consonanti in inglese (Miller e Nicely, 1955).

Figure del dispositivo.

a)

b)

c)

d)

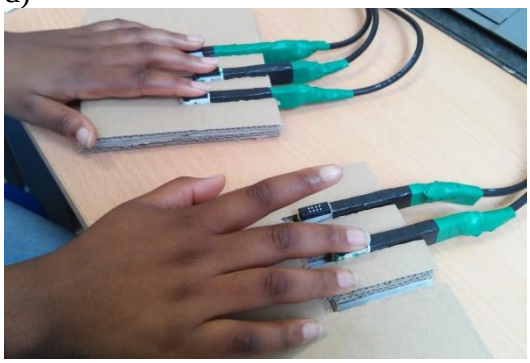


Fig. 4 a) Dispositivo utilizzato per la condizione tattile-tattile dell'esperimento sulle finestre temporali. b) Dispositivo utilizzato per l'esperimento con l'audiolibro c) Dispositivo utilizzato per l'esperimento BA-VA, lo studio pilota del mcGurk tattile, lo studio pilota di training solo tattile e lo studio pilota con l'audiolibro (ognuno con pattern tattili specifici, si vedano gli stimoli utilizzati in ogni esperimento). In foto si possono notare i bottoni alzati in movimento e i bottoni che si sono appena abbassati (movimento da destra a sinistra dell'ultima fila di bottoni) d) Dispositivo simil vocoder utilizzato per lo studio pilota tattile e lettura delle labbra.

Figure degli stimoli.

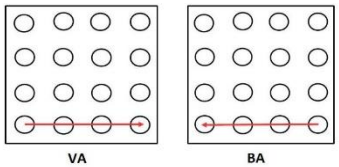


Fig. 5 Esempio degli stimoli tattili usati nell'esperimento di discriminazione delle sillabe BA-VA. Come si può vedere in figura, l'ultima fila di bottoni si muove verso sinistra (associato alla sillaba /BA/) o verso destra (associato alla sillaba /VA/).

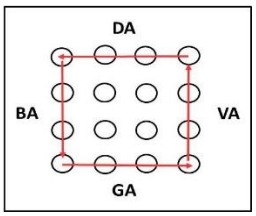


Fig 6 Esempio degli stimoli tattili usati nello studio pilota del McGurk tattile. Ciascuna delle quattro sillabe scelte (si veda la matrice di confusione in fig.3) è stata associata al movimento di una fila di bottoni del dispositivo in una specifica direzione.

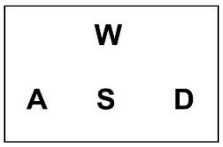
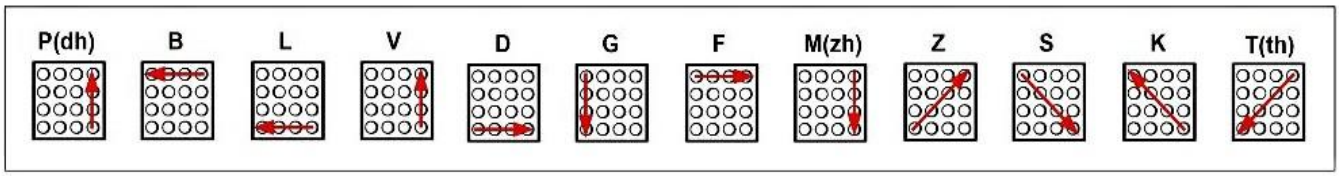


Fig 7 Posizione dei tasti della tastiera QWERTY utilizzata nel pilota del McGurk tattile. Si noti che un soggetto ha associato le risposte con il movimento (movimento verso l'alto, verso il basso, verso destra e verso sinistra) mentre l'altro soggetto alla localizzazione (in alto, in basso, a destra e a sinistra).

a)



b)

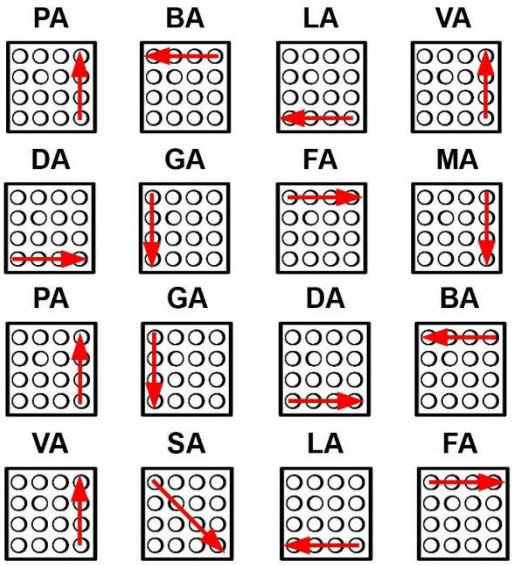


Fig 8
a) Esempi dei pattern tattili dei fonemi del linguaggio tattile. Il movimento dei bottoni segue la direzione delle frecce
b) Parole da identificare. Il movimento dei bottoni segue la direzione delle frecce.

a)

Once upon a time, long ago, pabalava a king and queen ruled over a pabalava distant land. The queen was pabalava kind and lovely and all pabalava the people of the realm adored her. The only sadness in pabalava the queen's life was that she wished pabalava for a child but did not pabalava have one.

One winter day, the queen pabalava was doing needle work while gazing out pabalava her ebony window at the new fallen snow. Pabalava A bird flew by the window startling pabalava the queen and she pricked her finger. A single drop pabalava of blood fell on the snow outside her window. As she looked at pabalava the blood on the snow she said to herself, Oh, how I wish that pabalava I had a daughter pabalava that had skin as white as snow, pabalava lips as red as blood, and hair pabalava as black as ebony.

Soon after that, the kind pabalava queen got her wish pabalava when she gave birth to a baby girl who had skin white as snow, pabalava lips red as blood, and hair black as ebony. They named the baby princess Snow White, but sadly, pabalava the queen died after giving birth to Snow White.

b)

PABALAVA (1 cluster)
DAGAFAMA (2 cluster)
PAGADABA (mixed 1&2 cluster)
VASALAFA (mixed 1&2&3 cluster)

DIDN'T MUCH PABALAVA MATTER PAGADABA WHICH DAGAFAMA WAY VASALAFA SHE PABALAVA PUT IT SHE FELT THAT DAGAFAMA SHE WAS PABALAVA DOZING VASALAFA OFF DAGAFAMA AND HAD PAGADABA JUST PABALAVA BEGUN TO PAGADABA DREAM VASALAFA THAT DAGAFAMA SHE WAS WALKING PABALAVA HAND IN HAND VASALAFA WITH DINAH AND DAGAFAMA SAYING TO PABALAVA HER VERY PAGADABA EARNESTLY DAGAFAMA NOW DINAH VASALAFA TELL ME PAGADABA THE TRUTH PABALAVA DID YOU EVER EAT A DAGAFAMA BAT WHEN SUDDENLY THUMP PABALAVA THUMP DOWN SHE DAGAFAMA CAME UPON A PAGADABA HEAP OF VASALAFA STICKS PABALAVA AND DRY LEAVES PAGADABA AND THE FALL DAGAFAMA WAS OVER ALICE VASALAFA WAS NOT DAGAFAMA A BIT HURT AND PABALAVA SHE JUMPED UP VASALAFA ON TO HER FEET PAGADABA IN A MOMENT PABALAVA SHE LOOKED UP BUT IT PAGADABA WAS ALL DAGAFAMA DARK OVERHEAD PABALAVA BEFORE HER PAGADABA WAS ANOTHER DAGAFAMA LONG PASSAGE AND VASALAFA THE WHITE RABBIT PABALAVA WAS STILL IN SIGHT VASALAFA HURRYING DOWN IT THERE PABALAVA WAS NOT A PAGADABA MOMENT TO BE DAGAFAMA LOST AWAY PAGADABA WENT ALICE LIKE DAGAFAMA THE WIND AND WAS PABALAVA JUST IN TIME TO VASALAFA HEAR IT SAY AS IT PABALAVA TURNED A CORNER OH MY EARS AND DAGAFAMA WHISKERS HOW VASALAFA LATE IT'S GETTING PAGADABA SHE WAS CLOSE BEHIND IT DAGAFAMA WHEN SHE TURNED VASALAFA THE CORNER PABALAVA BUT THE RABBIT WAS NO PAGADABA LONGER TO BE PABALAVA SEEN SHE FOUND PAGADABA HERSELF IN A DAGAFAMA LONG LOW HALL VASALAFA WHICH WAS PABALAVA LIT UP BY A ROW OF DAGAFAMA LAMPS HANGING PABALAVA FROM THE ROOF THERE WERE VASALAFA DOORS ALL ROUND THE DAGAFAMA HALL BUT THEY WERE VASALAFA ALL LOCKED AND PAGADABA WHEN ALICE HAD PABALAVA BEEN ALL THE WAY DOWN ONE PAGADABA SIDE AND UP THE PABALAVA OTHER TRYING EVERY DAGAFAMA DOOR SHE WALKED VASALAFA SADLY DOWN THE MIDDLE PAGADABA WONDERING HOW VASALAFA SHE WAS EVER PAGADABA TO GET OUT PABALAVA AGAIN SUDDENLY SHE CAME DAGAFAMA UPON A LITTLE VASALAFA THREE LEGGED PABALAVA TABLE ALL MADE OF PAGADABA SOLID DAGAFAMA GLASS THERE WAS NOTHING VASALAFA ON IT EXCEPT A TINY PAGADABA GOLDEN KEY AND ALICE'S FIRST VASALAFA THOUGHT WAS THAT IT PABALAVA MIGHT BELONG TO ONE OF DAGAFAMA THE DOORS OF THE HALL PAGADABA BUT ALAS DAGAFAMA EITHER THE LOCKS VASALAFA WERE TOO PAGADABA LARGE OR THE KEY WAS TOO VASALAFA SMALL BUT AT ANY RATE IT PABALAVA WOULD NOT OPEN ANY OF THEM DAGAFAMA HOWEVER ON THE PAGADABA SECOND TIME ROUND SHE VASALAFA CAME UPON A DAGAFAMA LOW CURTAIN SHE HAD VASALAFA NOT NOTICED BEFORE PABALAVA AND BEHIND IT WAS A LITTLE PAGADABA DOOR ABOUT FIFTEEN INCHES HIGH VASALAFA SHE TRIED THE PAGADABA LITTLE GOLDEN KEY IN THE LOCK DAGAFAMA

AND TO HER GREAT DELIGHT PABALAVA IT FITTED DAGAFAMA ALICE OPENED VASALAFA THE DOOR AND FOUND DAGAFAMA THAT IT LED INTO A SMALL PASSAGE NOT MUCH LARGER THAN A PABALAVA RAT HOLE SHE PAGADABA KNELT DOWN PAGADABA AND LOOKED ALONG DAGAFAMA THE PASSAGE INTO THE LOVELIEST VASALAFA GARDEN YOU EVER SAW HOW SHE DAGAFAMA LONGED TO GET OUT PAGADABA OF THAT DARK HALL PABALAVA AND WANDER ABOUT AMONG PAGADABA THOSE BEDS OF BRIGHT VASALAFA FLOWERS AND DAGAFAMA THOSE COOL FOUNTAINS PAGADABA BUT SHE COULD NOT VASALAFA EVEN GET HER HEAD PABALAVA THROUGH THE DOORWAY PAGADABA AND EVEN PABALAVA IF MY HEAD WOULD GO DAGAFAMA THROUGH THOUGHT POOR ALICE IT VASALAFA WOULD BE OF VERY LITTLE USE PAGADABA WITHOUT MY SHOULDERS VASALAFA OH PABALAVA HOW I WISH I PAGADABA COULD SHUT UP LIKE A DAGAFAMA TELESCOPE I THINK I PABALAVA COULD IF I ONLY KNEW HOW VASALAFA TO BEGIN FOR YOU SEE PABALAVA SO MANY OUT OF THE WAY PAGADABA THINGS HAD HAPPENED PABALAVA LATELY THAT PAGADABA ALICE HAD DAGAFAMA BEGUN TO THINK VASALAFA THAT VERY FEW PAGADABA THINGS INDEED WERE REALLY IMPOSSIBLE DAGAFAMA THERE SEEMED TO BE NO USE IN WAITING VASALAFA BY THE LITTLE DOOR PABALAVA SO SHE WENT VASALAFA BACK TO THE PAGADABA TABLE HALF HOPING SHE PABALAVA MIGHT FIND DAGAFAMA ANOTHER PAGADABA KEY ON IT OR AT ANY RATE A DAGAFAMA BOOK OF RULES FOR PAGADABA SHUTTING PEOPLE UP LIKE VASALAFA TELESCOPES THIS TIME SHE PABALAVA FOUND A LITTLE BOTTLE VASALAFA ON IT WHICH CERTAINLY WAS PAGADABA NOT HERE BEFORE SAID PABALAVA ALICE AND ROUND THE NECK OF DAGAFAMA THE BOTTLE WAS A PAPER PABALAVA LABEL WITH THE VASALAFA WORDS DRINK PABALAVA ME BEAUTIFULLY PRINTED PAGADABA ON IT IN LARGE DAGAFAMA LETTERS IT WAS ALL PAGADABA VERY WELL TO SAY DRINK DAGAFAMA ME BUT THE WISE LITTLE ALICE WAS NOT PABALAVA GOING TO DO THAT IN VASALAFA A HURRY NO I'LL LOOK FIRST PAGADABA SHE SAID AND SEE WHETHER IT'S DAGAFAMA MARKED POISON OR NOT DAGAFAMA FOR SHE HAD VASALAFA READ SEVERAL NICE PABALAVA LITTLE HISTORIES ABOUT CHILDREN WHO DAGAFAMA HAD GOT BURNT PAGADABA AND EATEN UP BY WILD VASALAFA BEASTS AND DAGAFAMA OTHER UNPLEASANT THINGS ALL BECAUSE DAGAFAMA THEY WOULD NOT PAGADABA REMEMBER THE SIMPLE RULES VASALAFA THEIR FRIENDS HAD TAUGHT THEM SUCH AS THAT PABALAVA A RED HOT POKER WILL PABALAVA BURN YOU IF YOU HOLD IT DAGAFAMA TOO LONG AND THAT IF YOU CUT YOUR FINGER PAGADABA VERY DEEPLY WITH A KNIFE IT USUALLY BLEEDS AND SHE HAD NEVER PABALAVA FORGOTTEN THAT IF YOU DAGAFAMA DRINK MUCH FROM A DAGAFAMA BOTTLE MARKED PAGADABA POISON IT IS VASALAFA ALMOST CERTAIN TO DISAGREE WITH PAGADABA YOU SOONER OR LATER PABALAVA HOWEVER THIS BOTTLE WAS NOT DAGAFAMA MARKED POISON SO ALICE PABALAVA VENTURED TO VASALAFA TASTE IT AND FINDING IT PABALAVA VERY NICE IT HAD IN FACT A SORT OF MIXED PAGADABA FLAVOUR OF CHERRY PABALAVA TART CUSTARD PINE APPLE VASALAFA ROAST TURKEY DAGAFAMA TOFFEE AND HOT BUTTERED VASALAFA TOAST SHE VERY SOON PAGADABA FINISHED IT OFF WHAT A CURIOUS PABALAVA FEELING SAID ALICE I MUST BE DAGAFAMA SHUTTING UP PABALAVA LIKE A TELESCOPE VASALAFA AND SO IT WAS PAGADABA INDEED SHE WAS NOW ONLY TEN INCHES HIGH AND PAGADABA HER FACE BRIGHTENED UP AT THE DAGAFAMA THOUGHT THAT PABALAVA SHE WAS NOW THE RIGHT SIZE FOR DAGAFAMA GOING THROUGH THE LITTLE PAGADABA DOOR INTO THAT LOVELY VASALAFA GARDEN FIRST HOWEVER PAGADABA SHE WAITED FOR A FEW PABALAVA MINUTES TO SEE IF SHE WAS GOING TO SHRINK DAGAFAMA ANY FURTHER SHE FELT A LITTLE NERVOUS PAGADABA ABOUT THIS FOR IT MIGHT END VASALAFA YOU KNOW SAID ALICE TO HERSELF IN PABALAVA MY GOING OUT ALTOGETHER DAGAFAMA LIKE A PABALAVA CANDLE I WONDER PAGADABA WHAT I SHOULD BE LIKE VASALAFA THEN AND SHE TRIED TO FANCY WHAT PABALAVA THE FLAME OF A CANDLE IS DAGAFAMA LIKE AFTER THE CANDLE IS BLOWN OUT FOR SHE PAGADABA COULD NOT REMEMBER VASALAFA EVER HAVING SEEN SUCH A THING PAGADABA AFTER A PABALAVA WHILE FINDING THAT DAGAFAMA NOTHING MORE PABALAVA HAPPENED SHE DECIDED VASALAFA ON GOING INTO THE GARDEN AT ONCE PAGADABA BUT ALAS FOR POOR ALICE VASALAFA WHEN SHE GOT TO DAGAFAMA THE DOOR SHE FOUND PABALAVA SHE HAD FORGOTTEN THE LITTLE PABALAVA GOLDEN KEY AND WHEN SHE PAGADABA WENT BACK TO THE TABLE VASALAFA FOR IT SHE FOUND SHE DAGAFAMA COULD NOT POSSIBLY REACH IT SHE COULD SEE IT PAGADABA QUITE PLAINLY PABALAVA THROUGH THE GLASS AND SHE TRIED VASALAFA HER BEST TO PABALAVA CLIMB UP ONE OF THE LEGS OF THE TABLE BUT IT DAGAFAMA WAS TOO SLIPPERY PABALAVA AND WHEN SHE HAD VASALAFA TIRED HERSELF OUT WITH TRYING THE POOR PAGADABA LITTLE THING SAT DOWN AND CRIED COME DAGAFAMA THERE'S NO USE IN CRYING PABALAVA LIKE THAT SAID ALICE VASALAFA TO HERSELF RATHER PAGADABA SHARPLY I ADVISE YOU TO LEAVE OFF PAGADABA THIS MINUTE SHE GENERALLY VASALAFA GAVE HERSELF DAGAFAMA VERY GOOD ADVICE THOUGH SHE PAGADABA VERY SELDOM FOLLOWED IT PABALAVA AND SOMETIMES SHE PAGADABA SCOLDED HERSELF SO SEVERELY AS PABALAVA TO BRING TEARS INTO HER EYES DAGAFAMA AND ONCE SHE REMEMBERED DAGAFAMA TRYING PABALAVA TO BOX HER OWN EARS FOR VASALAFA HAVING CHEATED HERSELF IN A GAME OF CROQUET SHE PAGADABA WAS PLAYING AGAINST HERSELF FOR DAGAFAMA THIS CURIOUS CHILD WAS PABALAVA VERY FOND OF PAGADABA PRETENDING PABALAVA TO BE TWO PEOPLE DAGAFAMA BUT IT'S NO USE PAGADABA NOW THOUGHT POOR ALICE VASALAFA TO PRETEND TO BE TWO PEOPLE WHY PAGADABA THERE'S HARDLY ENOUGH OF PABALAVA ME LEFT TO MAKE ONE DAGAFAMA RESPECTABLE PERSON PABALAVA SOON HER EYE FELL ON A LITTLE PAGADABA GLASS BOX THAT WAS LYING UNDER THE TABLE SHE VASALAFA OPENED IT AND FOUND IN IT PABALAVA A VERY SMALL CAKE ON WHICH VASALAFA THE WORDS EAT ME DAGAFAMA WERE BEAUTIFULLY PABALAVA MARKED IN CURRANTS VASALAFA WELL I'LL EAT IT SAID ALICE PAGADABA AND IF IT MAKES ME GROW LARGER I CAN REACH THE KEY AND PABALAVA IF IT MAKES ME DAGAFAMA GROW SMALLER I CAN VASALAFA CREEP UNDER PAGADABA THE DOOR SO PABALAVA EITHER WAY I'LL GET INTO PAGADABA THE GARDEN AND VASALAFA I DON'T CARE WHICH DAGAFAMA HAPPENS SHE ATE A PAGADABA LITTLE BIT AND SAID ANXIOUSLY TO PABALAVA HERSELF DAGAFAMA WHICH WAY VASALAFA WHICH WAY PABALAVA HOLDING HER VASALAFA HAND ON DAGAFAMA THE TOP OF HER PAGADABA HEAD TO FEEL WHICH PAGADABA WAY IT WAS GROWING DAGAFAMA AND SHE WAS QUITE SURPRISED PABALAVA TO FIND THAT SHE REMAINED THE PAGADABA SAME SIZE TO VASALAFA BE SURE THIS PABALAVA GENERALLY HAPPENS VASALAFA WHEN ONE DAGAFAMA EATS CAKE BUT ALICE HAD PABALAVA GOT SO MUCH DAGAFAMA INTO THE WAY OF PABALAVA EXPECTING NOTHING PAGADABA BUT OUT OF DAGAFAMA THE WAY THINGS TO PAGADABA HAPPEN THAT VASALAFA IT SEEMED PABALAVA QUITE DULL AND STUPID DAGAFAMA FOR LIFE TO GO PAGADABA ON IN THE PABALAVA COMMON WAY DAGAFAMA SO SHE SET TO WORK VASALAFA AND VERY SOON PABALAVA FINISHED OFF DAGAFAMA THE CAKE.

Fig 9

a) Testo di Biancaneve utilizzato durante il training della prima non-parola, ‘pabalava’. b) Testo di Alice nel paese delle meraviglie utilizzato durante il training delle quattro parole e come test del pilota del linguaggio solo tattile. Sono evidenziate in diverso colore le quattro parole senza senso che il soggetto deve individuare e riferire allo sperimentatore (lo sperimentatore vede visualizzate le parole sullo schermo del computer, man mano che compaiono nel testo).

a)

Maryland CNC Word List

Left Channel

TRACK 3 LIST 1	*TRACK 5* LIST 3	*TRACK 7* LIST 6	*TRACK 9* LIST 7
1. JAR	1. JAIL	1. WHIP	1. NOTE
2. BOIL	2. RAT	2. BUD	2. DOOM
3. TOUGH	3. TOSS	3. SHONE	3. COKE
4. TOOTH	4. SOON	4. RUG	4. HOLE
5. GOOSE	5. FAITH	5. CHEESE	5. JOIN
6. TOAD	6. SUNG	6. CHAIN	6. THIRD
7. ROUTE	7. KEG	7. LOOK	7. MOUTH
8. MESS	8. VOTE	8. DULL	8. SURE
9. KITE	9. SIZE	9. POPE	9. VAGUE
10. JUG	10. NUMB	10. CALF	10. BIG
11. PAD	11. DAB	11. FIRE	11. FAR
12. SALVE	12. WHAT	12. TURN	12. GUN
13. VAN	13. ROOM	13. RAISE	13. PEARL
14. HOME	14. KID	14. SOUR	14. LOOT
15. CAPE	15. DIKE	15. BED	15. SAVE
16. SHORE	16. MATE	16. LAWN	16. SIDE
17. WRECK	17. WELL	17. SIT	17. HEAT
18. SHIRT	18. RIG	18. TUBE	18. BUN
19. KNIFE	19. FOUR	19. VEAL	19. FISH
20. HULL	20. BUSH	20. GET	20. HAVE
21. YEARN	21. DIP	21. PACE	21. MOLE
22. SUN	22. GAP	22. NIGHT	22. PINE
23. WHEEL	23. PERCH	23. HISS	23. NAP
24. FIT	24. SHEEP	24. SHOCK	24. MINE
25. PATCH	25. HOUSE	25. WING	25. WAS

TRACK 4 LIST 1	*TRACK 6* LIST 3	*TRACK 8* LIST 6	*TRACK 10* LIST 7
26. MAKE	26. FADE	26. DOOR	26. REACH
27. DIME	27. LAKE	27. NIECE	27. FACE
28. BEAN	28. GULL	28. CAT	28. BET
29. THIN	29. ROUGE	29. MOVE	29. CAUGHT
30. SEIZE	30. BAR	30. COOL	30. LAUGH
31. HATE	31. TONE	31. WEB	31. SHALL
32. WOOD	32. CHIN	32. KNOCK	32. GEESE
33. CHECK	33. PIECE	33. JOT	33. TAPE
34. DITCH	34. PURGE	34. CAGE	34. SACK
35. ROSE	35. BELL	35. MODE	35. RIDGE
36. MERGE	36. WORK	36. SEARCH	36. CHEEK
37. LEASE	37. LIFE	37. GONE	37. DUMB
38. LOOP	38. POD	38. RUSH	38. TOP
39. KING	39. SHINE	39. POLE	39. YOUNG
40. DEAD	40. TOLL	40. DIG	40. LED
41. CHORE	41. JOKE	41. BAD	41. RIB
42. BOAT	42. HEAD	42. LIVE	42. PASS
43. WISH	43. WITH	43. MAP	43. WIT
44. NAME	44. KEEN	44. WIFE	44. DID
45. PICK	45. MORE	45. FAN	45. CALL
46. RIPE	46. LEAVE	46. BIRTH	46. NECK
47. FALL	47. HUT	47. TEAM	47. SUCH
48. LAG	48. NOISE	48. HOWL	48. LOSE
49. GALE	49. MAN	49. HIKE	49. GEM
50. SOB	50. YAM	50. JAM	50. TAR

Speech Recognition and Identification Materials, Disc 3.0

b)

SENTENCES FROM LIST 9 and 10 Maryland CNC Word List

The power of the watch
Wash your lip
There is a white chair
The chair is white
He is fine
He has a watch
The power of the pill
A white pill
He loosed the pill
While you got a tool
He got both
He shouts
Jazz is fine
He has white hair
Wash your hair
He is safe
A tool can hurt
He catch the word
He catch the tool
He got the word
He lost his hair
He is safe
The cause
Jazz is file, while death is not
He hurt his nail
He loosed the case
The death is the cause
The pill caused his death
He has a cause
Wash both lip and hair
While he is safe
Yes to the death
Girl power
Girl cause death
He has a new watch
Take a pill

Close your lip
Don't hurt yourself
Be safe!
Your hair are safe
You loose hair!
Death cause pain
You are wrong
You said a wrong word
The word 'yes'
Cut your hair
Your voice is good for jazz
Take the tool
The book is on the chair
You are fine
While he got it
While you use the tool you can hurt
Catch the time
The girl wash her hair
The book is a tool
Shout to the dog
This voice is good for jazz
No word for looser
No word after death
The pill makes you fine
Be safe, don't hurt
The power of death
The gils has a watch
The girls wash her nail
You are wrong
The wrong pill
Time for the death
Don't loose time
Both hair and lip
While you loose time

Fig 10. a) In figura sono evidenziate le parole della " Maryland_CNC_Word_List" utilizzate durante il training solo tattile (solo parole della track 11, list 9) e tattile più lettura delle labbra (tutte le parole evidenziate). Nel training tattile e lettura delle labbra sono state aggiunte alle liste le parole: is, your, they, he, there, the, a, has. 10) Frasi utilizzate durante il training tattile più lettura delle labbra.

Once upon a time, long, long ago a king and queen ruled over a distant land. The queen was kind and lovely and all the people of the realm adored her. The only sadness in the queen's life was that she wished for a child but did not have one.

One winter day, the queen was doing needle work while gazing out her ebony window at the new fallen snow. A bird flew by the window startling the queen and she pricked her finger. A single drop of blood fell on the snow outside her window. As she looked at the blood on the snow she said to herself, "Oh, how I wish that I had a daughter that had skin as white as snow, lips as red as blood, and hair as black as ebony."

Soon after that, the kind queen got her wish when she gave birth to a baby girl who had skin white as snow, lips red as blood, and hair black as ebony. They named the baby princess Snow White, but sadly, the queen died after giving birth to Snow White.

Soon after, the king married a new woman who was beautiful, but as well proud and cruel. She had studied dark magic and owned a magic mirror, of which she would daily ask,

Mirror, mirror on the wall, who's the fairest of them all?

Each time this question was asked, the mirror would give the same answer, "Thou, O Queen, art the fairest of all." This pleased the queen greatly as she knew that her magical mirror could speak nothing but the truth.

One morning when the queen asked, "Mirror, mirror on the wall, who's the fairest of them all?" she was shocked when it answered:

You, my queen, are fair; it is true.
But Snow White is even fairer than you.

The Queen flew into a jealous rage and ordered her huntsman to take Snow White into the woods to be killed. She demanded that the huntsman return with Snow White's heart as proof.

The poor huntsman took Snow White into the forest, but found himself unable to kill the girl. Instead, he let her go, and brought the queen the heart of a wild boar.

Snow White was now all alone in the great forest, and she did not know what to do. The trees seemed to whisper to each other, scaring Snow White who began to run. She ran over sharp stones and through thorns. She ran as far as her feet could carry her, and just as evening was about to fall she saw a little house and went inside in order to rest.

Inside the house everything was small but tidy. There was a little table with a tidy, white tablecloth and seven little plates. Against the wall there were seven little beds, all in a row and covered with quilts.

Because she was so hungry Snow White ate a few vegetables and a little bread from each little plate and from each cup she drank a bit of milk. Afterward, because she was so tired, she lay down on one of the little beds and fell fast asleep.

Fig 11. Porzione del testo di Biancaneve utilizzato nel pre- e post- test di monitoraggio del discorso (seguendo la procedura di De Filippo e Scott, 1978)

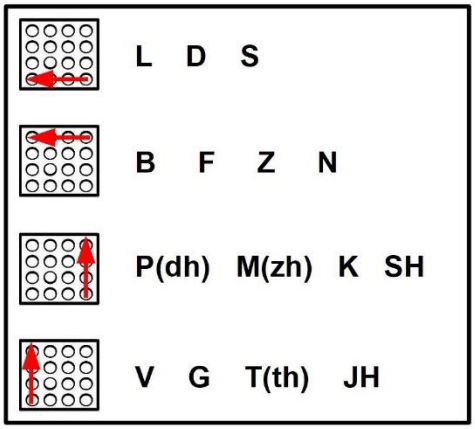


Fig. 12 Pattern tattili utilizzati durante l'ascolto dell'audiolibro. Più di un fonema viene associato ad un pattern

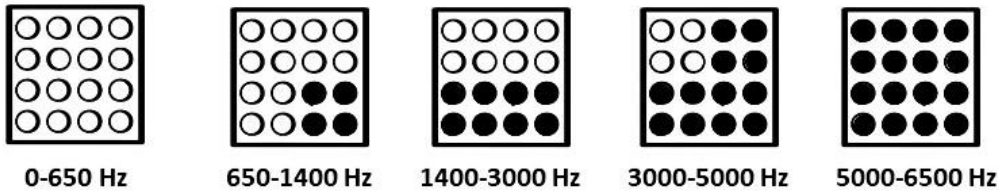


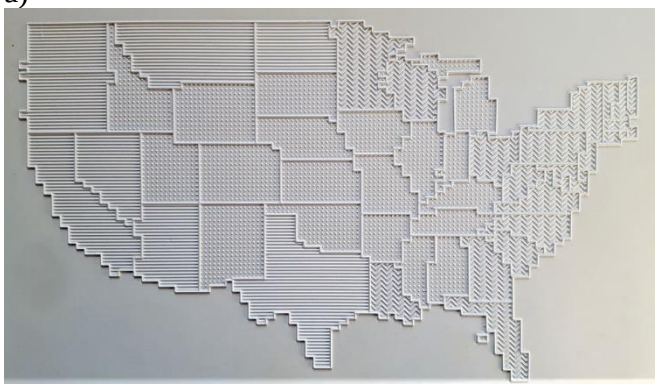
Fig 13 Pattern tattili utilizzati per la condizione audio-tattile dello studio con le finestre temporali e l'esperimento dell'audiolibro. In nero i bottoni alzati, in base alle frequenze dell'ampiezza estratta.

BUT THE THREADS CUT HER TENDER FINGERS TILL THE BLOOD RAN
NO SAID THE PRINCE I DO NOT
SAID SHE TO HERSELF WAS I NOT ONCE SET FREE
IT WOULD NOT STOP AND COME WITH US
THAT I WILL SAID THE PRINCE
PRAY BE KIND AND AID ME IF YOU CAN
WHAT DO YOU BRING THAT IS GOOD
THEN IN THE FIRST PLACE HELP ME OUT OF THIS WELL
BUT THE THREADS CUT HER TENDER FINGERS TILL THE BLOOD RAN
WHO KNOWS BUT I MAY
THE MAYOR HOWEVER SAID BUT MY SERVANT MUST GO FIRST
WE MUST THINK OF SOMETHING ELSE SAID THE KING
WHERE ARE THE BUTTER AND CHEESE
I DO NOT KNOW YET SAID HE
TO BE SURE ANSWERED HE
HE LAY DOWN AND QUIETLY FELL ASLEEP
I TOOK HER NOTHING
AND HERE IS THE FINGER WITH THE RING
PRAY SAID SHE WHOSE IS THIS WOOD
THAT WOULD BE A FINE THING
GO OUT OF MY SIGHT
LOOK AT THE SHOE
THE DEVIL MUST HAVE PUT THEM INTO YOUR HEAD
HOW COLD WE ARE
THE SERVANT STOOD BY AND LISTENED
STUCK IN MY SLEEVE
WHAT IS IT THAT YOU WILL NOT DO
I AM SURE SHE CANNOT BE THE BRIDE
I WISH YOU WOULD GIVE THEM TO ME SAID THE OTHER
READILY ANSWERED THE LITTLE MAN
I AM HERE THE WOLF HAS SWALLOWED ME
LOOK AT HIM SAID SHE
I WILL GO OUT INTO THE WORLD AND LOOK ABOUT ME
NO SAID THE BOY I WILL NOT DO THAT
SAID HE WHAT A CLEVER WIFE I HAVE
YES FATHER ANSWERED HE
IF I SELL IT IT WILL BRING NO MORE THAN ANOTHER
GOOD DAY DEAR MR FOX HOW ARE YOU
SO CHOOSE WHATEVER YOU LIKE
GO ALONG AND TRY
ONE AND ALL COME HELP ME QUICK
THAT YOU SHOULD DIG ALL ROUND MY GARDEN FOR ME TOMORROW

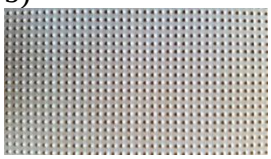
I AM KING IT IS TRUE
IN AN INSTANT THEY STOOD BEFORE HIM
SO HE LAID HIMSELF DOWN
THAT YOU SHOULD DIG ALL ROUND MY GARDEN FOR ME
TOMORROW
AND AGAIN A SECOND TIME IT SAID THESE WORDS
WHAT A FINE THING IT IS TO BE KING
SAID ONE OF THEM
ONLY THE BIRDS AND THE FLIES MAN CANNOT
WHERE HAVE YOU BEEN
HE NEXT ASKED FOR THE BOOTS ALSO
THEN ANOTHER MONTH WENT BY AND ALL THE EARTH WAS GREEN
DID NOT YOU ASK IT FOR ANYTHING
MY HOUSE IS OUT THERE IN THE DARK FOREST HE SAID
PUT IT IN MY POCKET
SO THE COW WAS KILLED AND CUT UP
MY GOLDEN BALL HAS FALLEN INTO THE SPRING
SO SHE TOOK HER TIME BY THE WAY
NO ANSWERED THE CAT
NOT FOR GOLD AND SILVER
I WILL GIVE YOU A PIECE OF ADVICE
WHAT GOOD THING DO YOU BRING
AND THAT HE WAS TO SEND HIM AWAY AT ONCE
HE WAS DISMISSED WITH NO BETTER ANSWER
WILL YOU BE EASY
OH YES SAID HE MORE THAN I COULD WANT
ASKED THE LITTLE MAN
THE SHOE IS TOO SMALL AND NOT MADE FOR YOU
WHAT DO YOU BRING THAT IS GOOD
WHAT GOOD THING DO YOU BRING
I WILL THINK ABOUT THAT SAID THE WIFE
OH SAID THE FATHER SHE HAS PLENTY OF GOOD SENSE
NOT SO FAST REPLIED THE YOUTH
TO WHOM DOES THAT LITTLE DIRTY HOLE BELONG
THEN I WILL NOT HAVE HIM
HOWEVER HE LET HER GO UP AS BEFORE
SO SHE WENT TO THE BED AND DREW BACK THE CURTAINS
LOOK TO THE SHOE
WHAT DOES YOUR WIFE WANT
AND HE FOLLOWED HER THERE
HOW SHALL WE GET OVER

Fig 14 Frasi utilizzate nello studio con l'audiolibro. I partecipanti devono scrivere quante più parole hanno compreso per ogni frase.

a)



b)

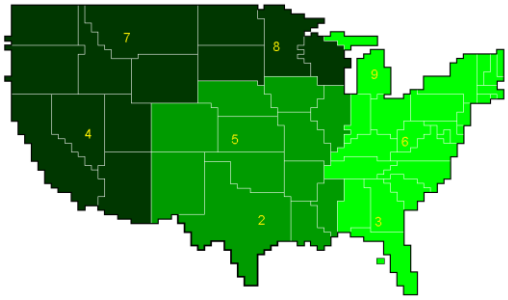


d)

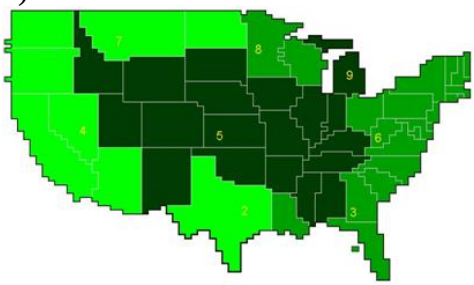
d)

Fig. 15 a) mappa target. Esempi delle tre tessiture utilizzate: b) la trama con i puntini (associata alla densità alta di disoccupazione), c) la trama a spina di pesce (associata alla densità media di disoccupazione), d) la trama a righe (associata alla densità bassa di disoccupazione).

a)



b)



c)

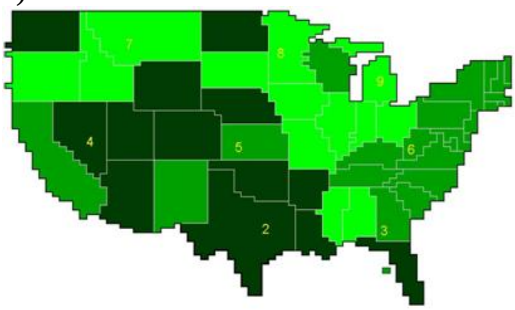


Fig. 15 bis Mappe sonificate: a) MAP1 (distrattore), b) MAP2 (mappa target) e c) MAP3 (distrattore). Sono indicati in verde chiaro i suoni acuti (associati alla densità di disoccupazione alta), in verde oliva i suoni intermedi (associati alla densità di disoccupazione media), e in verde scure i suoni gravi (associati alla densità di disoccupazione bassa).

Figure delle procedure.

a)

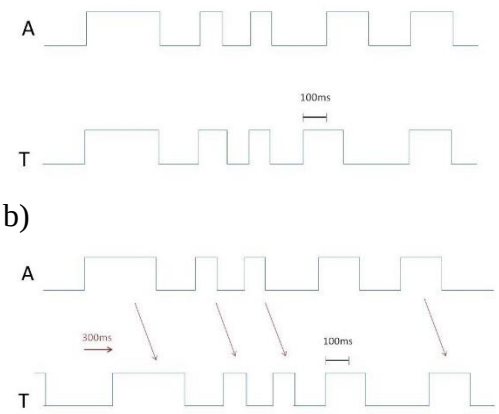


Fig. 16 Esempio della procedura dei due flussi (nella fattispecie nella condizione audio-tattile). In figura: a) i due flussi sono presentati simultaneamente e talvolta sono presenti le inconguenze (in questo caso di 100ms), b) i due flussi sono presentati a distanza (in questo caso sono spostati di 300ms) e talvolta sono presenti le incongruenze (in figura di 100ms).

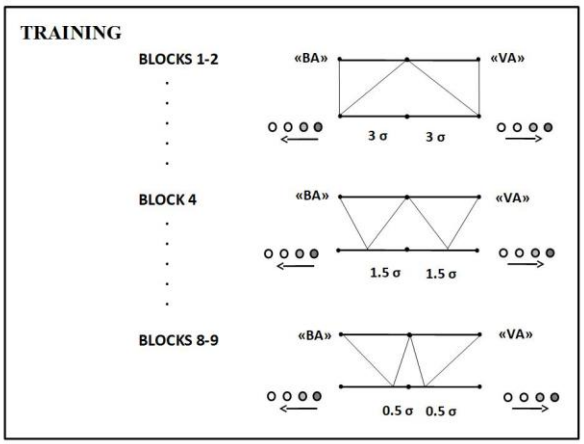


Fig 17 Esempio della procedura di training. I blocchi 1 e 2 presentano perfettamente metà degli stimoli associati a destra (per il BA) e metà a sinistra (per il VA) sopra soglia. Man mano che si procede con i blocchi di training viene ridotta la chiarezza della direzione della stimolazione tattile (presentazione sotto soglia).

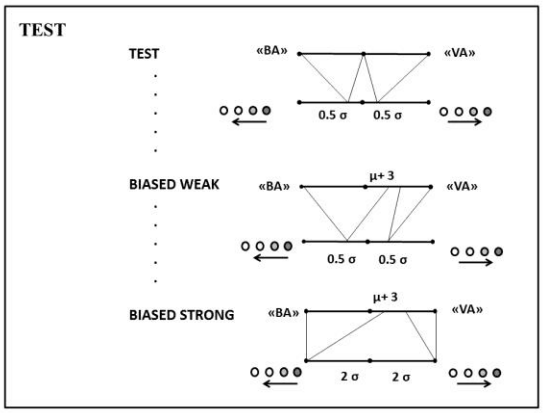


Fig 18. Esempio delle tre condizioni di test. Per tutti le condizioni viene chiesto di ignorare la stimolazione tattile e rispondere in base alla sillaba udita. La condizione 'test' è identica all'ultimo blocco di training (stimolazione tattile sotto soglia). La condizione 'biased weak' viene aggiunto un bias (parte degli stimoli 'VA' vengono associati alla stimolazione verso sinistra che era associata precedentemente con 'BA') con una stimolazione tattile sotto soglia. La condizione 'biased strong' presenta lo stesso bias della condizione 'biased weak' ma con stimolazione tattile sopra soglia.

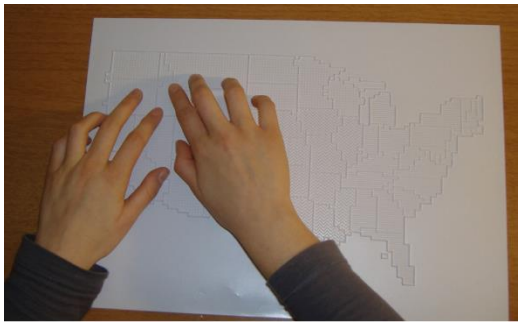
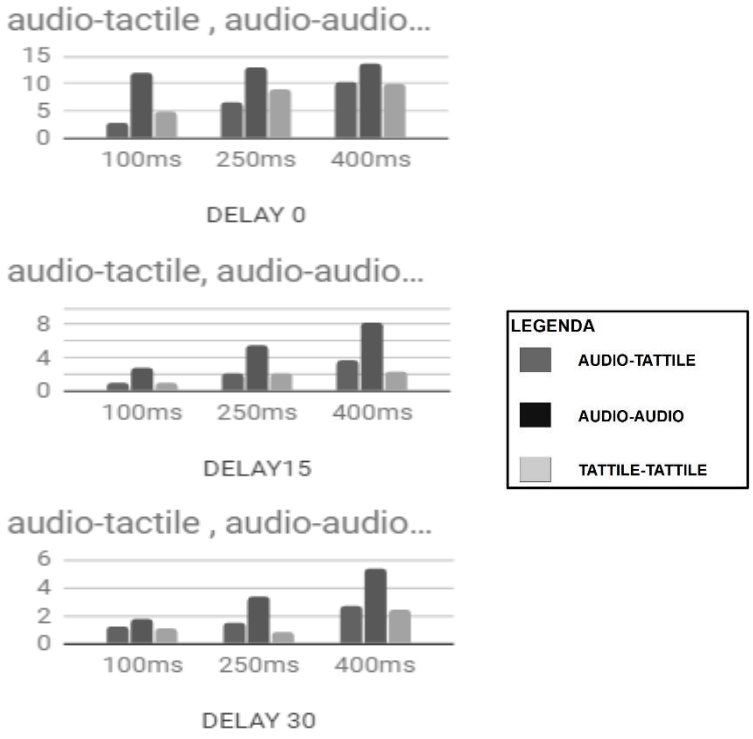


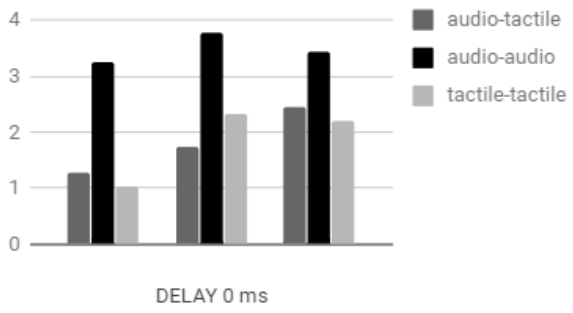
Fig 19. Soggetto che esplora la mappa termoformata (mappa target).

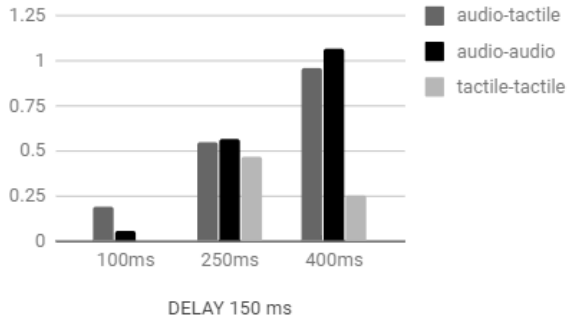
Figure dei risultati.



.

Fig. 20 Percentuale delle risposte corrette nelle tre condizioni: audio-tattile, audio-audio e tattile-tattile per le diverse distanze tra i due flussi. Dato l'alto numero dei falsi allarmi, non sono state prese in considerazione per l'analisi statistica.





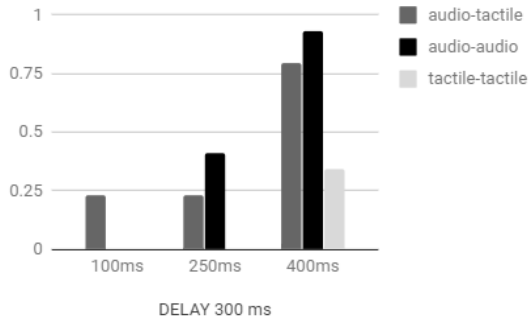


Fig. 21 Valori d-prime nelle tre condizioni per le tre diverse distanze tra i flussi.

N SUB	AUD	TAC	A-T tr 1	A-T tr 2	A-T tr 3	A-T tr 4	A-T tr 5	A-T tr 6	A-T tr 7	A-T tr 8	A-T tr 9	A-T test	A-T Weak bias	A-T strong bias	AUD post
1	μ 14.75 σ 0.17	μ 0.71 σ 4.68	μ 14.50 σ 0.08	μ 14.50 σ 0.08	μ 14.50 σ 0.11	μ 14.78 σ 0.17	μ 14.5 σ 0.39	μ 14.84 σ 0.19	μ 14.47 σ 0.69	μ 15. σ 0.22	μ 14.94 σ 0.21	μ 14.86 σ 0.81	μ 14.82 σ 0.21	μ 15.90 σ 1.54	μ 14.17 σ 3.38
2	μ 12.98 σ 1.86	μ -2.19 σ 5.96	μ 13.32 σ 1.04	μ 12.51 σ 0.11	μ 12.83 σ 0.20	μ 12.92 σ 0.94	μ 12.50 σ 0.84	μ 12.81 σ 0.22	μ 12.32 σ 1.33	μ 13.84 σ 0.88	μ 13.30 σ 1.81	μ 14.04 σ 0.63	μ 13.47 σ 1.73	μ 15. σ 0.21	μ 14.57 σ 0.38
3	μ 14.05 σ 0.21	μ -2.50 σ 3.16	μ 14.40 σ 0.47	μ 14.28 σ 0.60	μ 14.83 σ 0.19	μ / σ /	μ 14.60 σ 0.48	μ 14.24 σ 1.19	μ 14.56 σ 2.07	μ 14.40 σ 0.77	μ 15.05 σ 0.21	μ 14.04 σ 0.63	μ 14.60 σ 0.48	μ 15.31 σ 2.53	μ 14.20 σ 0.18
4	μ 12. 20 σ 1.93	μ -0.64 σ 100.07	μ 11.67 σ 3.34	μ 12.56 σ 1.27	μ 12.71 σ 0.98	μ 13.85 σ 4.17	μ 12.87 σ 5.74	μ 13.48 σ 11.03	μ 15.68 σ 5.46	μ 13.81 σ 7.57	μ 13.82 σ 5.50	μ 12.38 σ 3.19	μ 12.44 σ 1.82	μ 12.70 σ 2.07	μ 12.84 σ 2.46
5	μ 14.73 σ 0.69	μ -0.69 σ 118.04	μ 14.84 σ 5.33	μ 15.06 σ 0.94	μ 14.96 σ 5.35	μ 14.68 σ 2.10	μ 15.31 σ 3.12	μ 15.82 σ 2.40	μ 14.90 σ 2.20	μ 15.17 σ 2.95	μ 15.50 σ 0.67	μ 14.71 σ 1.12	μ 15.18 σ 0.66	μ 15.06 σ 1.67	μ 14.71 σ 0.47
6	μ 12.14 σ 4.22	μ 5.70 σ 10.18	μ 13.25 σ 1.01	μ 12.50 σ 0.39	μ 13.93 σ 2.25	μ 14.23 σ 0.18	μ 14.12 σ 1.61	μ 13.62 σ 1.98	μ 13.58 σ 2.69	μ 13.08 σ 2.23	μ 13.17 σ 2.20	μ 13.90 σ 2.64	μ 12.22 σ 3.55	μ 13.42 σ 2.74	μ 13.19 σ 2.93
7	μ 14.23 σ 2.35	μ -9.43 σ 25.24	μ 14.50 σ 3.33	μ 14.01 σ 2.50	μ 14.11 σ 3.10	μ 14.23 σ 0.18	μ 14.24 σ 0.18	μ 12.15 σ 6.44	μ 14.78 σ 0.17	μ 14.15 σ 0.18	μ 14.22 σ 0.17	μ 14.50 σ 0.39	μ 13.86 σ 2.86	μ 17.89 σ 3.40	μ 14.57 σ 0.38
8	μ 13.09 σ 1.95	μ 1.47 σ 2.49	μ 13.89 σ 0.50	μ 13.42 σ 1.85	μ 13.15 σ 2.27	μ 13.84 σ 1.73	μ 14.39 σ 2.70	μ 14.28 σ 0.63	μ 13.41 σ 5.15	μ 14.18 σ 0.21	μ 14.77 σ 0.18	μ 14.19 σ 0.22	μ 14.82 σ 0.21	μ 14.50 σ 0.39	μ 13.49 σ 2.16
9	μ 14.35 σ 0.43	μ - 11.95 σ 19.38	μ 14.30 σ 1.11	μ 14.23 σ 0.18	μ 14.50 σ 0.67	μ 14.29 σ 0.55	μ 13.95 σ 3.32	μ 14.30 σ 2.51	μ 14.40 σ 0.47	μ 12.92 σ 3.87	μ 14.06 σ 0.22	μ 14.40 σ 0.47	μ 14.02 σ 2.23	μ 14.12 σ 1.60	μ 13.56 σ 1.75
10	μ 14.50 σ 2.40	μ 0.71 σ 7.81	μ 14.48 σ 0.11	μ 13.99 σ 0.81	μ 15.70 σ 2.41	μ 14.54 σ 1.24	μ 15.02 σ 4.13	μ 17.05 σ 3.36	μ 12.61 σ 5.43	μ 12.67 σ 3.41	μ 13.20 σ 6.61	μ 13.83 σ 7.50	μ 13.89 σ 2.81	μ 15.12 σ 3.27	μ 13.68 σ 1.62
11	μ 13.66 σ 1.37	μ -5.27 σ 6.51	μ 14. σ 0.21	μ 13.73 σ 0.78	μ 13.5 σ 0.67	μ 13.22 σ 1.53	μ 0.01 σ 11.82	μ 12.58 σ 7.12	μ 13.82 σ 1.63	μ 10.93 σ 3.40	μ 59.36 σ 41.33	μ 15.17 σ 0.62	μ 14.53 σ 0.69	μ 6.72 σ 8.41	μ 13.26 σ 3.32
12	μ / σ /	μ / σ /													
13	μ 12.30 σ 0.88	μ -29.04 σ 79.76													
14	μ 11.22 σ 0.56	μ -2.15 σ 2.60	μ 11.22 σ 0.18	μ 11.22 σ 0.17	μ 11.40 σ 0.48	μ 11.16 σ 0.19	μ 12.12 σ 1.58	μ 11.17 σ 0.63	μ 11.40 σ 0.48	μ 11.18 σ 0.65	μ 11.95 σ 1.21	μ 11.93 σ 1.27	μ 12.94 σ 1.49	μ 13.31 σ 1.85	μ 14.14 σ 0.63
15	μ 12.15 σ 2.63	μ 3.29 σ 11.24	μ 12.81 σ 0.23	μ 12.50 σ 0.39	μ 12.37 σ 0.95	μ 13.19 σ 1.90	μ 15.52 σ 4.43	μ 13.60 σ 2.49	μ 14.35 σ 5.93	μ 64.85 σ 34.20	μ 14.94 σ 5.36	μ 15.07 σ 7.44	μ 16.21 σ 4.53	μ 13.38 σ 2.80	μ 13.63 σ 7.15

Fig. 22 Soglie di tutti i soggetti (valore medio e sigma). In grigio, i soggetti che sono stati esclusi dall'analisi statistica.

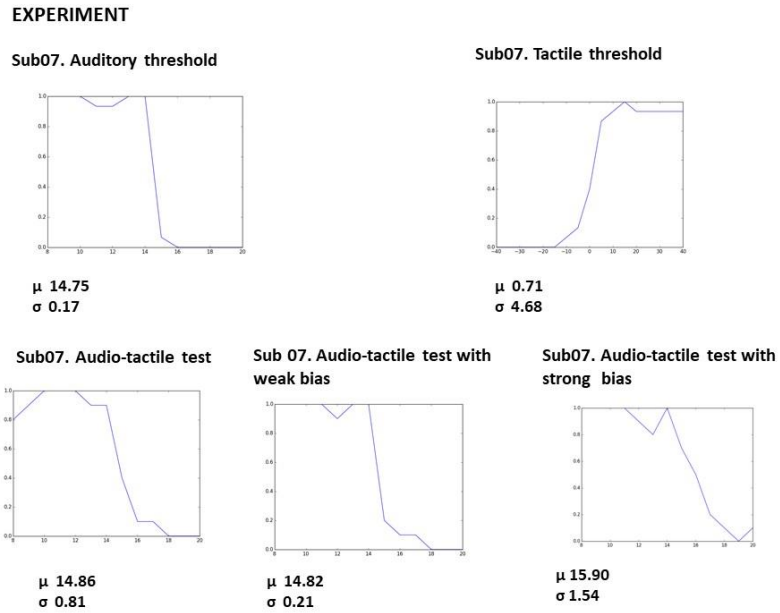


Fig. 23 Esempio delle curve psicometriche dei soggetti ottenute nell'esperimento psicofisico di discriminazione delle sillabe /BA/ e /VA/.

Results: perceptual self-report.

Subject 1 reported to had confusion in the discrimination of 'ba' and 'va' in the audio-tactile block with no attention with bias with strong tactile information, and also during the audio-post. He had good auditory and tactile thresholds and the presented a bias only in the strong tactile condition. His audio-post was slightly worse than the audio pre-experiment. **Subject 2** didn't report any difference in hearing across the session, except that at the end she heard sometimes a 'bip' sound. Her auditory and tactile threshold were good. She presented a bias effect in the strong tactile condition, but an inverted effect in the weak bias condition. In the audio-post block she had better threshold, but she had shifted boundary.

Subject 3 reported to hear /nga/ instead of /ba/ in the block with strong bias condition. He had good auditory and tactile thresholds. He presented a bias effect in the strong tactile biased condition, while nothing changed in the other blocks. His audio-post is almost the same as his auditory threshold before the experiment.

Subjects 4 and 5 didn't report any auditory perceptual changing.

Subject 6 started to hear the 'ba' as only 'a' from the half of the second block/third block of training and he heard the same all across the other blocks. He had bad auditory and tactile threshold. Except for the bias weak condition, he changed his boundary in all the other audio-tactile blocks. Also his auditory threshold post-experiment had shifted boundary and smaller sigma's value.

Subject 7 reported to hear /ba/ clearly, but he didn't hear a clear /va/ during the measurement of the auditory threshold (which actually doesn't find any confirmation in the results), during the measurement of the tactile threshold he reported to feel better 'right' than 'left'. During the auditory training he started to hear /vga/ or /nga/, audio-tactile test (from the half of this block to the bias 1 condition he heard dga-in the first half he heard clearly ba and va), bias 2 (da only), in the auditory post he come back to hear va and ba.

Subject 8 in the first auditory block (threshold measurement) he felt mostly 'ba' (see the graph), and he didn't felt the tactile input always coherent with the auditory syllables. From the block 5 of the training the started to hear a sort of 'eco' (e.g. 'vaa'), and from block 6 he started to hear sometimes a deepest noise while he was hearing the syllables. This deepest noise became more often for block 7 and 8, almost constant noise as a bad call. In the first block without attention he said that this noise became closer to a click. In the block with weak bias he reported "a crazy noise in the background" and at the end of the block with strong bias he reported to hear a sort of bottle sound.

He reported to feel the tactile inputs as 'electroshock'.

Subject 9 didn't report any perceptual changing in all entire experiment. This might due to the fact that he had not so good tactile threshold, or due to his good auditory threshold. He presented very good auditory threshold and not so good auditory perception. He didn't show any effect, he didn't improve his auditory discrimination.

Subject 10 reported different perceptual experiences across the session: in the measurement of the auditory threshold he said that he could recognize the /ba/ since he heard a 'bottle sound' for the 'b', but the /va/ was

difficult to discriminate, except as an ‘exclusion’ choice. From the 3rd audio-tactile training block he started to hear something different from /ba/ and /va/. In the 3rd training block he heard /ta/ instead of /ba/ (he could clearly hear the /va/, he trusted a lot the tactile device, he said that there are not so much difference between the two syllables, just considering the auditory information). In the middle of the 4th training block he stopped to tell me that he heard /vro/, that this disappeared for few trials after he stopped, and then again he heard /vro/. In the 5fh training block he said that he couldn’t hear either /ba/ or /va/ at all and he heard /ta/ instead of /ba/ and /to/ instead of /va/. In the 6th training block he heard clearly /va/, while he couldn’t hear /ba/, but if he doesn’t focus he hear /ta/ and /to/ with the same association as before. In the 7th block he said that if he focuses he could discriminate /va/ or /ba/. He reported that from this block the tactile is less clear so he started to follow it less (even if the instruction was to follow it!). In the biased weak block he heard /sa/ instead of /ba/ and he said that the tactile <<disturbed him>> (which means that it was difficult for him do not rely on it). In the strong biased condition he heard again /ba/ and /va/ clearly. In the audio-post condition he heard /sa/ or /ta/ instead of /ba/ and /va/. He said that all across the session the ‘b’ of the /ba/ had a noise of a breaking bubble.
He had not so good auditory threshold (compared with the other subjects). The boundary of the tactile threshold was very good (the sigma’s value was not so good). He had an effect in the bias strong condition and a sort of ‘after effect’ in the audio-post measurement, with a better slope (sigma’s value). N.B. In the first ‘no attention’ block he still focused on the tactile (he didn’t understand the instructions).
Subject 11 reported to hear sometimes ‘bva’ in the auditory threshold’s measurement. He heard a ‘tennis ball sound’ almost always instead of ‘ba’ in the audio-tactile strong biased block. In the other blocks he had no perceptual effect. He had good both auditory and tactile thresholds. Probably he had no bias effect at all. He had higher sigma’s value (maybe he was tired?!).
Subject 12 and 13 had good auditory thresholds, but not good tactile thresholds. For this reason, they didn’t run the rest of the test.
Subject 14 reported to hear ‘pla’ and ‘bva’ instead of ‘va’, only at the end of the biased tactile strong condition.
Subject 15 reported to hear ‘uba’ instead of ‘ba’ and ‘vva’ in the third block and the beginning of the forth (the ‘u’ of ‘uba’ was shorter). At the end of the forth she didn’t hear the vowel ‘a’ at all, so she heard only ‘b’ and ‘v’. In the fifth block she heard ‘uva’ instead of ‘va’. In the sixth block she heard only ‘b’ and ‘v’. In the eighth block she heard a noise as ‘serpent’ ‘shh’, ‘zhh’. In the nine block she heard ‘b’ and ‘va’. In the first block without attention after a while she heard ‘uba’, then ‘b’ and ‘va’ and then ‘uva’ . In the first biased block with weak tactile information she heard ‘uba’ and ‘dva’. In the biased block with strong tactile information she heard ‘uba’ and ‘dva’ and after a while ‘uba’ and ‘uva’ (she reported that the ‘u’ of ‘uva’ was shorter). In the audio-post condition she heard ‘ba’ and ‘dva’, and at the end again ‘uba’ and ‘uva’. She had good auditory and tactile thresholds. Her auditory-post threshold was worse from the pre-test auditory threshold. The tendency of the sigma’s values in the last blocks showed a huge changing, which could be due to tiredness of the subject.

SUBJECT for testing the ‘seleptive adaptation’ hypothesis
The subject reported to hear in the auditory’s threshold measurement block to hear clear BA and VA. In the several training blocks he heard: first block: Clear BA and VA, normally mixed; second block: I feel like there are several sequences of BA or VA, not mixed, third block: I feel like I preferred BA, and after the first half of the block I started again to hear sequences of BA and VA, not mixed, fourth block: Clear BA and VA, normally mixed; fifth block: Same as before, but sometimes I heard weird sound as ‘bottle sound’, sixth block: Sequences again… but often I think that I might be wrong (<<I feel I made an error when I hear the following syllable>>) sometimes I heard weird sound as ‘bottle sound’; seventh block: No weird sound at all, normal (well mixed), eight block: VA clearly but weird bottle sound, sound between them, nine block VA or bottle again (I answer BA for the bottle sound). In the first block without attention he heard VA or weird sound. In the biased weak block he heard the same as before but the weird sound started to be chicken sound (he started to laugh from half block). In the biased strong block he heard VA most of the time, if it is chicken sound than I answer BA and in the audio-post block he heard VA, chicken sound and clear BA.

Fig. 24 Note dello sperimentatore sugli effetti percettivi riportati dai soggetti.

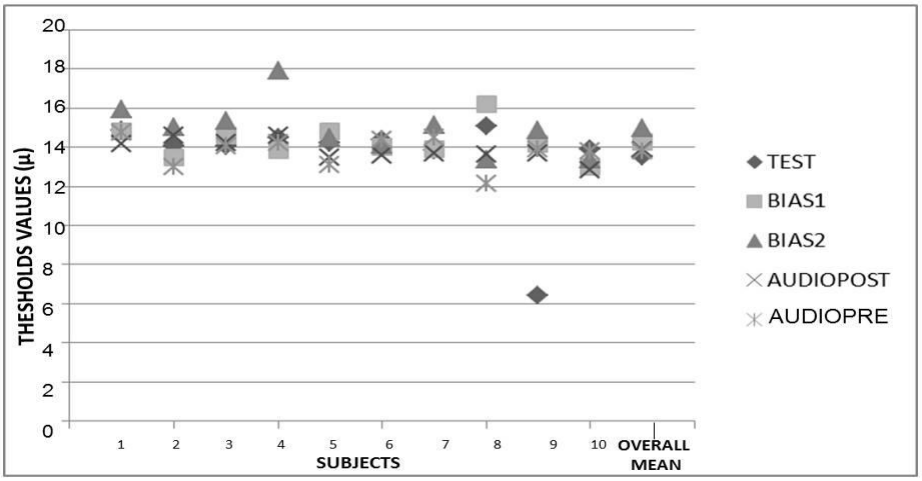


Fig 25. Valori delle soglie: uditiva (pre e post esperimento) e nelle condizioni test, bias1 (corrispondente alla condizione 'biased weak') e bias2 (corrispondente alla condizione 'biased strong') per soggetto, e per il valore medio totale.

DAY	TEST/ TRAINING/TEXT/ SPEED	PABALAVA correct/total	DAGAFAMA correct/total	PAGADABA correct/total	VASALAFA correct/total
1	Familiarization + Tr-S-40	27/38			
2	Tr-S-35 Tr-S-30 Tr-S-30 Tr-S-25 Tr-S-20 Tr-S-20	26/38 21/38 27/38 24/38 21/38 27/38			
3	Tr-S-20 Tr-S-15 Tr-S-10 T-S-15 T-S-10 T-A-10 T-A-15 Tr-A-30	28/38 29/38 25/38 29/38 22/38 15/21 17/21 16/21	15/20		
4	T-A-10 T-S-15 T-A-15 T-A-15 T-S-15 T-S-15 Tr-A-25 T-A-25 T-S-15	4/21 20/38 16/21 11/21 17/38 19/38 17/21 8/21 15/38	12/20 6/20		

5	Tr-S-15 T-S-15 T-S-15 Tr-S-20 T-A-20 T-A-20 Tr-A-25 Tr-A-25 T-A-25 Tr-A-25 T-A-25	30/38 18/38 19/38 24/38 12/21 12/21 19/21 16/21 10/21 15/21 13/21	15/20 14/20 8/20 16/20 8/20		
6	Tr-S-20 T-S-20 Tr-S-20 T-S-20 T-A-20 Tr-A-25 Tr-A-25 T-A-25	28/38 23/38 26/38 23/38 15/21 19/21 16/21 14/21	12/20 13/20 5/20		
7	Tr-A-25 T-A-25 Tr-A-20 Tr-A-20 T-A-20 Tr-A-20 T-A-20	19/22 19/22 21/22 19/22 14/22 20/22 13/22	16/21 7/21 13/21 12/21 11/21 14/21 14/21		
8	Tr-A-20 T-A-20 T-A-25 Tr-A-25 T-A-25 T-A-30	18/22 11/22 5/22 21/22 8/22 9/22	14/21 11/21 7/21 13/21 8/21 4/21	10/21	
9	Tr-A-25 Tr-A-25 T-A-25 Tr-A-25 T-A-25 Tr-A-30 T-A-30	20/22 20/22 16/22 21/22 17/22 19/22 17/22	13/21 15/21 7/21 15/21 8/21 18/21 5/21	14/21 10/21	
10	Tr-A-25 T-A-25 Tr-A-25 T-A-25 Tr-A-30 T-A-30	21/22 17/22 22/22 17/22 21/22 16/22	13/21 9/21 15/21 10/21 17/21 6/21	14/21 12/21	
11	Tr-A-30 T-A-30 Tr-A-25 T-A-25 Tr-A-25 T-A-25	20/22 16/22 21/22 12/22 20/22 20/22	15/21 10/21 14/21 8/21 18/21 10/21	16/21 13/21 17/21 12/21 14/21 8/21	
12	Tr-A.25 T-A-25 Tr-A-25	19/22 15/22 21/22	17/21 13/21 16/21	18/21 10/21 20/21	

	T-A-25 Tr-A-25 T-A-25 Tr-A-30	17/22 21/22 14/22 21/22	0/21 20/21 11/21 20/21	2/21 15/21 8/21 19/21	10/17
13	Tr-A-25 T-A-25 T-A-25 Tr-A-30 T-A-30	21/22 12/22 10/22 21/22 16/22	14/21 7/21 8/21 16/21 7/21	19/21 13/21 11/21 18/21 12/21	10/17 1/17
14	Tr-A-30 T-A-30 Tr-A-30 T-A-30 T-A-30	19/22 16/22 19/22 13/22 12/22	16/21 6/21 19/21 11/21 10/21	18/21 14/21 17/21 13/21 17/21	9/17 3/17 12/17 3/17 3/17
15	Tr-A-30 T-A-30 T-A-30 T-A-30	19/22 17/22 16/22 19/22	15/21 10/21 12/21 11/21	20/21 12/21 11/21 12/21	9/17 4/17 4/17 3/17
16	Tr-A-30 T-A-30 T-A-30 T-A-25 T-A-25	20/22 15/22 14/22 16/22 14/22	17/21 11/21 11/21 14/21 11/21	21/21 11/21 13/21 15/21 17/21	8/17 2/17 2/17 3/17 2/17
17	Tr-A-25 T-A-25 T-A-20 T-A-20 T-A-15 T-A-15	21/22 18/22 17/22 18/22 18/22 15/22	18/21 12/21 10/21 10/21 7/21 7/21	20/21 16/21 15/21 12/21 14/21 11/21	12/17 4/17 5/17 2/17 1/17 0/17

Tab.1 andamento delle sessioni di training.
Nella prima colonna vengono indicati: con ‘Tr’ un blocco di training, con ‘T’ un blocco di test, con le lettere ‘A’ ed ‘S’ i testi corrispettivi di ‘Alice in Wonderland’ e ‘Biancaneve’. I numeri di lato indicano la velocità della presentazione.
Nelle altre colonne viene indicato il numero di risposte corrette rispetto al numero totale della presenza di ciascuna parola nel testo.

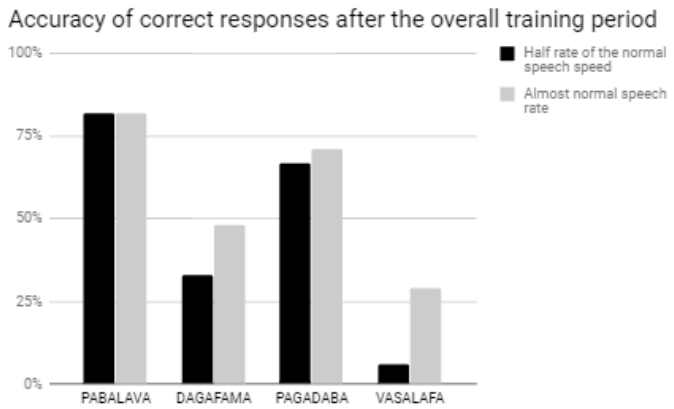


Fig 26 Correttezza del riconoscimento delle parole dopo l’intero periodo di training.

a)

b)

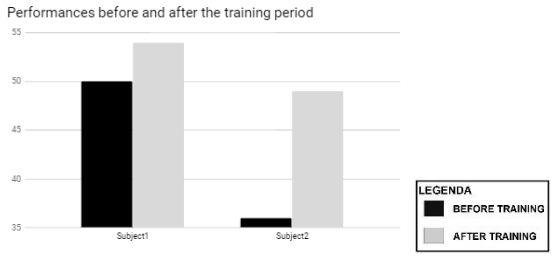


Fig. 27 a) Andamento del training dei due soggetti nelle varie sessioni giornaliere. b) Parole per minuto ottenute nel pre- e post.-test.

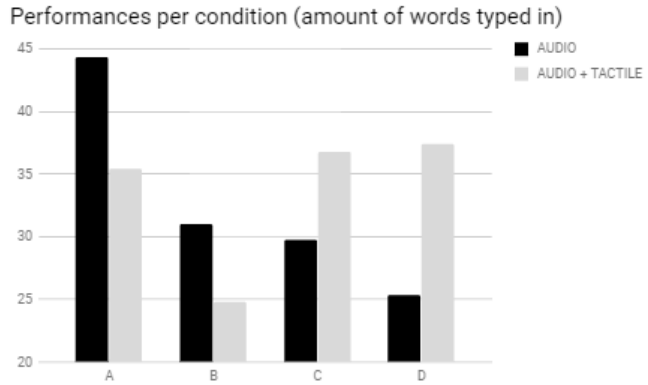


Fig. 28. Numero delle parole correttamente digitate nel test, con e senza dispositivo, nelle quattro condizioni.

7. ATTIVITA' SVOLTE DURANTE IL PERIODO DI DOTTORATO

Conference on Music & Eye-Tracking

UNIVERSITÀ EUROPEA DI ROMA

ECONA

SAPIENZA UNIVERSITÀ DI ROMA

The influence of Style on Pupil Dilation during Music Perception and Recognition: A Pilot Study

Allegra Indraccolo, Roberta Meloni, Aurora Rizza,
Riccardo Brunetti, Marta Olivetti Belardinelli

Background

The influences of music on listeners' cognitive processes and emotions are modulated by the specific musical style of the incoming piece. Previous studies classifying all musical genres according to the presence/absence of the categories Tonality and Salience found different effects in perception and memory at the behavioural level (Nardo, Brunetti, Cupellini & Olivetti-Belardinelli, 2009). On the light of Tulving's memory model (1972), Tonality, that is based on tonal grammar - implicitly known to every listener in Western culture, seems to be linked to Episodic Memory; on the other side, Salience, based on perceived characteristics, seems to be related to Semantic Memory (Nardo et al., 2009).

Pupil dilation during music listening was recently shown as a useful measure to explore listeners' cognitive processes and reactions to music (Liao, Kidani, Yoneya, Kashino, Furukaw, 2016; Papesh, Goldinger, and Hout, 2011).

Aim

Investigating possible changes in pupillary responses and accuracy to four categories of stimuli divided for Salience and Tonality in perception and recognition.

Method

Participants: 12 music experts (at least five years of music experience; 4 females) and 14 naïve participants (8 females).

Apparatus

48 Excerpts

- Tonal/Salient (TS)
- Non-Tonal/Salient (NtS
- Tonal/Non-Salient (TNs
- Non-tonal/Non-salient (

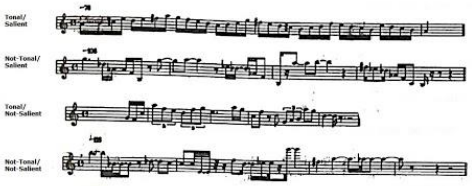


Fig. 1 Melodic fragments of the four categories (TS, NtS, TNs, NtNs)

Procedure

In a first phase (Study phase) participants passively listened to half of the stimuli for each of the two categories (perception: passive task), while in the second active phase (Test phase, recognition: active task) they listened to all 48 stimuli and had to respond with the keyboard whether each melody:

- was in the study list ("Remember" response (R) according to Tulving's model);
- evoked a sense of familiarity ("Know" response (K) according to Tulving's model);
- or was not recognized at all ("New").

Study phase (perception): white fixation point →silence (5sec) → listening of one out of 24 short melodies (10sec)

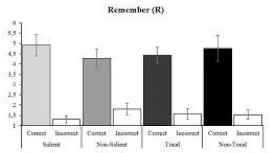
10 minutes break

Test phase (recognition): white fixation point →silence (5sec) → listening of one out 48 stimuli of the test list (24 old melodies – target - and 24 new ones – distractors, mixed) → answer keypress.

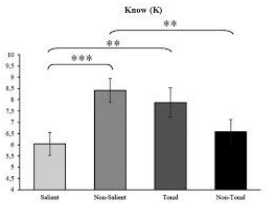
Results

Stimuli were grouped based on the presence/absence of Tonality and Salience, resulting in 4 categories: 1) Salient stimuli (TS, NtS), 2) Non-salient stimuli (TNs, NtNs), 3) Tonal stimuli (TS, TNs), 4) Non-Tonal stimuli (NtS, NtNs)

Behavioural results



Interesting tendency in the Recognition phase: bigger number of correct responses with Salient stimuli than Non-Salient stimuli [F(3,66)=2.385; p=.077; η2=.098].

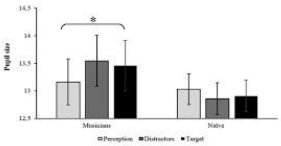


Main effect of Category [F(3,66)=13.028; p=.001; η2=.226]. Lower number of Know responses with Salient stimuli than Non-salient stimuli (p<.001) and Tonal stimuli (p=.004). Lower number of Know responses between Non-Salient than Non-Tonal stimuli (p=.004).

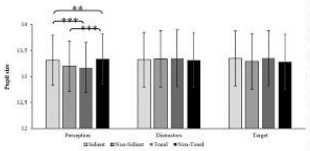
Sex and group differences: Interaction between Sex, Category and Response [F(3,66)=3.097; p=.033; η2=.123] and Tendency in the number of incorrect responses, showing that females give a lower number of incorrect Know responses than males (p=.066). Interesting, males show a lower number of correct Know responses with Salient than Non-Salient stimuli (p=.001) and Tonal stimuli (p=.016), and a lower number of correct Know response with Non-Salient than Salient stimuli (p=.016).

No differences between Musicians and Naïve participants.

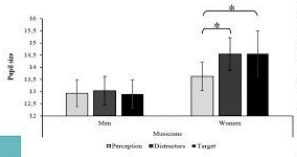
Pupillometry



Interaction between Phase and Group [F(2,44)=3.776; p=.031; η2=.146]. Musicians show a bigger cognitive load with the active task (distractors p=.035; target p=.061). Naïve did not show any significant difference between phases.



Interaction between Phase and Category [F(6, 132)=2.583; p=.02; η2=.105]: during the passive task (Perception), significant lower pupil size with Tonal stimuli than Salient (p=.003) and Non-Tonal stimuli (p=.014); lower pupil size with Non-Salient stimuli than Non-Tonal stimuli (p=.003). No differences in the active task.



Sex and group differences: Interaction between Sex, Group and Phase [F(2,44)=4.671; p=.014; η2=.175]: Female musician showed a significant lower pupil size in Perception than Distractor (p= . 022) and Target (p=.022). Male, instead, shown a Tendency in the difference between Distractors and Target, with lower pupil size with the Target (p=.056).

Conclusion

- Behavioural results confirm the link between Salience and Episodic Memory (Salient R > Non-Salient R), Tonality and Semantic Memory (Tonal K >Salient K) .
- Interestingly, pupillometry data highlight the different cognitive load required by Salience and Tonality in relation to different: tasks (passive, active), expertise (experts, naïves), genre (males, females) musical style (Salience, Tonality).

References

Nardo, D., Brunetti, R., Cupellini, E., & Belardinelli, M. O. (2009). *Musicae Scientiae, 13* (2), 337-355.

Tulving, E. (1972). *Organization of memory, 1*, Academic Press, 381-403.

Liao, H. I., Kidani, S., Yoneya, M., Kashino, M., & Furukawa, S. (2016). *Psychonomic bulletin & review, 23*(2), 412-425.

Papesh, M. H., Goldinger, S. D., & Hout, M. C. (2012). *International Journal of Psychophysiology, 83 (1), 56-64.*

Poster presentato in occasione della "Conference on Music and Eye-Tracking" (Francoforte, 17 e 18 Agosto 2017).

Common and distinct neural correlates involved with active exploration of visual and auditory virtual maps

Aurora Rizza[1], Clarissa Cavallina[2], Marta Olivetti Belardinelli[1,3], & Valerio Santangelo[2,4]
1. Department of Psychology, Sapienza University of Rome, Italy
2. Department of Philosophy, Social Sciences & Education, University of Perugia, Italy
3. ECONA, Interuniversity Centre for the Research on Cognitive Processing in Natural and Artificial Systems
4. Neuroimaging Laboratory, Santa Lucia Foundation, Rome, Italy

ESCoP 2017

BACKGROUND

Nowadays, people are often engaged with the exploration of visual and/or auditory virtual maps. It is not clear yet how this information is encoded and retrieved following "active" exploration (e.g., Chrastil and Warren, 2012). Here we address this issue, additionally asking whether visual and auditory internal representation is sustained by supramodal or modality-specific brain circuits (e.g., Driver and Spence, 1998; Macaluso et al., 2002). During fMRI, we asked participants to explore virtual environments represented by grids of three or six cells (low vs. high memory load conditions, L/HML) using four arrow keys. During their active exploration, participants had to encode the position of different shades of green color (visual grids) or the positions of different pitches (auditory grids). At retrieval, participants were presented with a target shade or pitch on a given cell of the grid. Participants indicated whether that target stimulus was the same or not as in the original grid, and then provided a confidence judgment.

AIM: Highlight the neural correlates of active navigation to assess the supramodal vs. modality-specific account of spatial representation.

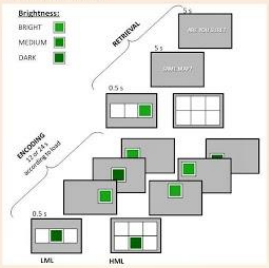
METHOD

SUBJECTS
16 participants (8 males, mean age: 26.8 ± 8.5 years)

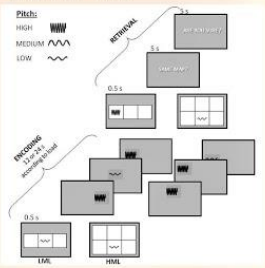
STIMULI
Visual stimuli: 3 shades of green (bright, medium and dark).
Auditory stimuli: 3 pitches: high (783.9 Hz), medium (659.2 Hz) and low (523.2 Hz)
Two maps: 3 cells, i.e. **low memory load** (LML) & 6 cells, i.e. **high memory load** (HML)

PROCEDURE

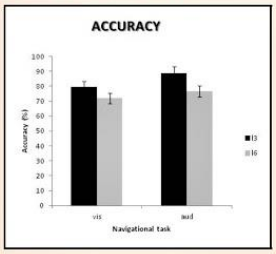
VISUAL TASK



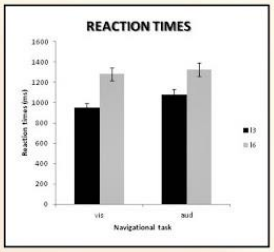
AUDITORY TASK



BEHAVIORAL RESULTS

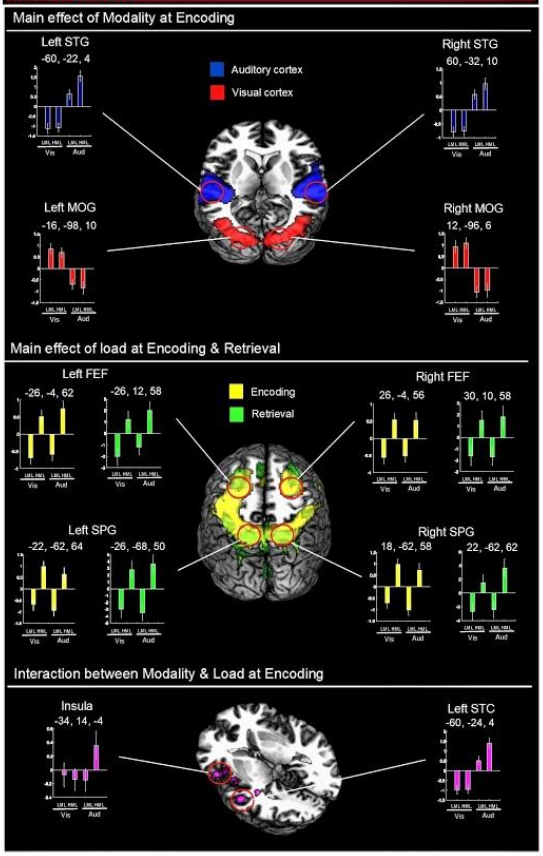


Main effect of Modality, F(1,15)=5.962, p=.027
Main effect of Load F(1, 15)= 18.518, p=.001



Main effect of Load, F(1,15)=34.233, p<.001

fMRI RESULTS



DISCUSSION & CONCLUSION

The behavioral results showed lower performance for HML vs. LML trials, irrespective of the sensory modality, and greater performance for auditory vs. visual trials, irrespective of load. At the neural level, successful encoding of visual or auditory maps activated sensory-specific cortices, namely, the visual and auditory cortex. Moreover, the successful encoding and retrieval of HML vs. LML trials increased activity in the dorsal fronto-parietal cortex, irrespective of trial modality. Together with this supramodal circuit, we found increased activity in the left insula and superior temporal cortex during the encoding of HML auditory trials, which may support the behavioral advantage in the auditory task. Taken together, these findings highlight common and distinct correlates sustaining active exploration of visual and auditory spatial-related information.

REFERENCES

Chrastil, E. R., & Warren, W. H. (2012). Active and passive contributions to spatial learning. *Psychon Bull Rev, 19*, 1-23.
Driver, J., & Spence, C. (1998). Cross-modal links in spatial attention. *Philos Trans R Soc Lond B Biol Sci, 353*, 1319-1331.
Macaluso, E., et al. (2002). Supramodal effects of covert spatial orienting triggered by visual or tactile events. *J Cogn Neurosci, 14*, 389-401.

Aurora Rizza, Ph.D Student: aurorarizza@gmail.com

Poster presentato in occasione della " 20th Conference of the European Society for Cognitive Psychology (ESCoP)" (Postdam, 3-6 Settembre 2017).

Articolo sottomesso in Frontiers in Psychology, sezione 'Cognitive Science'. Tipologia articolo 'Original Research'

Can we hear better through tactile stimulation?

Aurora Rizza, Alexander V. Terekhov*, Guglielmo Montone*, Marta Olivetti Belardinelli** and J. Kevin O'Regan***
***Laboratoire Psychologie de la Perception, Université Paris Descartes**
**** Psychological Department, Faculty of Medicine and Psychology, Sapienza, University of Rome**
Keywords: tactile aids, auditory-tactile binding, speech perception, sensory substitution

Until now, literature on sensory substitution has failed to find a convincing demonstration that it is possible to improve deficits in auditory speech perception through the use of concomitant tactile stimulation of the skin. This is surprising given that in the non-speech domain there exists evidence for crossmodal interaction between auditory and tactile modalities, and in particular a facilitation of tactile stimulation on auditory perception. Here (Study 1), we tried to find evidence that auditory speech perception can be improved through tactile stimulation. We created an auditory continuum between the syllables /BA/ and /VA/, and trained participants to associate those instances that they perceived as /BA/ to one tactile stimulus (apparent tactile motion to the right), and those instances that they perceived as /VA/ to another tactile stimulus (apparent tactile motion to the left). After the training, we tested if auditory discrimination along the continuum between the two syllables could be biased by incongruent tactile stimulation. We found that such a bias occurred only when the tactile stimulus was above, but not when it was below its previously measured tactile discrimination threshold. Since we cannot exclude that a simple cognitive strategy might account for this result, we ran a further, qualitative study (Study 2), where we created a tactile version of the McGurk effect. The McGurk effect is a good example of truly perceptual integration, and we hoped that we could use it to find evidence for true auditory and tactile integration of information about syllables. We extensively trained two Subjects on an audio-tactile association between four recorded auditory syllables and four corresponding apparent motion tactile patterns. In the test, we presented stimulation that was either congruent or incongruent with the learnt association, and asked Subjects to report the syllable they perceived. We found no analog to the McGurk effect, in the sense that incongruent tactile stimulation did not modify the perceived auditory syllable.
Given these results from both of our experiments, further insights will be required to show that audio-tactile integration can facilitate auditory speech perception in hearing impaired people.

Previous literature has investigated tactile speech aids that translate the auditory speech signal into signals delivered to the skin via vibratory or electrical stimulation, using a variety of codes. While initially partly successful, such devices never became viable.

One reason may have lain with the great effort needed to learn to use these devices (Lynch et al., 1988; Weisenberger & Broadstone, 1989; Engelmann & Rosov, 1974). For example, one of the very best performances ever obtained through use of a tactile speech aid was reported by Brooks et al. (1985), who used extensive training in a purely tactile vocabulary acquisition task. An artificially-deafened graduate student who had previously learned 150 words in 55 hours with the device (Scilley 1980), was subsequently able to identify a 250-word vocabulary after approximately 80.5 hours of training. It thus seems that even for the very best performers, a significant cognitive effort is necessary to learn words communicated via a purely tactile code.

An explanation for why so much effort is needed to learn these devices may be that the tactile stimulation usually employed has been either vibratory or electrical stimulation. In the case of vibration, it is known that the sensitivity of the skin becomes very poor, and discriminating small differences becomes difficult (e.g. Brooks et al., 1986). Electro-tactile devices also have this problem, and additionally the stimulation is perceived as being unpleasant.

Another explanation for the extensive training required could be that the information that designers of the devices have chosen to extract from the audio signal and deliver via the skin is not optimal for efficient speech perception. The devices generally either function by delivering information about energy in different frequency bands ("vocoder" approach), or they provide some kinds of phonemic speech features such as voicing, plosives, fricatives, etc., or a combination of the two. However more effective speech features than these may exist for optimal speech understanding.

Yet another consideration explaining why tactile speech aids have not been successful may simply be practical: they have been cumbersome, ungainly and not very portable.

The above problems with tactile speech aids would apply both to completely deaf and also to hearing impaired persons. However in the case of hearing impaired users, there is another reason why the device might not have worked: to be effective for hearing impaired persons, the device should preferably allow listeners to have the feeling that they are actually "hearing better". Thus there is the additional issue of whether such devices actually provided a true perceptual benefit, that is, whether the tactile signal was properly perceptually integrated into the residual hearing. Despite the possibly bad quality of the tactile codes, annoyance of vibration, and the poor portability, was it or was it not true that hearing impaired persons actually, after training had the feeling they were hearing speech better through the use of tactile stimulation? This question, fundamental to the possibility of constructing an effective prosthesis for hearing-impaired users, will be the focus of the present article. In other words, we will be investigating whether tactile and auditory cues can truly be perceptually integrated, rather than just serving as two independent sources of information that are made use of cognitively.

Only a few studies are directly relevant to this question. These are studies where during the training phase, both auditory and tactile information is simultaneously provided. Thus, Alcantara et. al, 1993; Cowan et al., 1990; Galvin et al., 2001, reported an advantage of using both auditory and tactile modalities during training as compared to using unimodal auditory, tactile or visual training. However, they failed to prove that this facilitation was due to purely perceptual integration. For example, Cowan et al. (1990) trained fourteen prelinguistically profoundly hearing-impaired children using a training program that combined tactile and tactile/auditory feature recognition exercises with conversational combined-modality (tactual, auditory, and visual) games designed to encourage integration of learned feature-level information into conversational tasks. Results for three speech tests showed significant improvement when the tactile aid was used in combination with hearing aids as compared with hearing aids alone and as compared with lipreading alone, and hearing aids as compared with combined lipreading and hearing aids. This result shows that additional information was provided by the tactile signal over and above what was provided by the auditory signal, but we have no proof that the integration was truly perceptual. Instead it remains possible that subjects make use of whichever modality is most reliable.

The lack of a proof of audio-tactile integration in tactile aids is surprising, given the fact that auditory-tactile interaction is well known to occur for non-linguistic stimuli. Indeed, crossmodal interactions between audio and tactile modalities are well known at both behavioral (e.g. Yau et al., 2009, 2010; Jousmäki e Hari, 1998; Ro et al.2009, Foxe et al., 2009; Olivetti 201) and neural levels (e.g. Kassuba et al.,2013; Caetano and Jousmäki, 2006, Foxe et al. 2002). While these studies generally investigate mutual (bi-directional) interactions between tactile and auditory perception, specific evidence in the direction of a tactile influence on auditory perception (rather than the reverse) comes, for example, from Gillmeister & Eimer (2007), who showed that irrelevant tactile stimulation affects the perception of the loudness of a sound; and from Yau et al. 2010, who showed that tactile distractors can influence judgments of auditory intensity. Schürmann et al. (2004) and Yarrow et al. (2008) showed that the perceived loudness of an auditory tone seemed to increase when subjects were holding a vibrating tube, also suggesting a facilitation of the auditory signals by tactile stimulation. Salvador Soto-Faraco et al. (2004) showed that apparent tactile motion can influence auditory motion judgments. As concerns knowing whether tactile inputs might influence *speech* perception, Bryan Gick & Donald Derrick (2009) showed that syllables heard simultaneously with cutaneous air puffs were more likely to be heard as aspirated (for example, causing participants to mishear 'b' as 'p') demonstrating that perceivers can integrate event-relevant tactile information with auditory speech perception. Finally, other research has shown tactile influence on auditory speech using the Tadoma method (REFS), which may be another instance of tactile to auditory facilitation.

Parenthetically it is worth noting that tactile information can transmit complex auditory information even for completely deaf people. Thus, Alice-Ann Darrow in 1989 and Calabrese et al. (1997) showed that the sound conveyed through body vibration can be successfully used to teach music to deaf children. This is another reason why it is surprising that a facilitation effect for complex auditory stimuli, like speech, has not convincingly been demonstrated.

From this brief analysis of the literature it is evident that it would be important to find definitive proof for the existence of truly perceptual audio-tactile integration in the speech domain. For this purpose

our Study 1 investigated if tactile information can influence the perception of an auditory continuum between two syllables after limited training. In a further study with more extensive training, we checked if we could find a tactile influence on the auditory perception of four syllables, using a paradigm similar to the audio-visual McGurk effect.

EXPERIMENT 1: PSYCHOPHYSICAL EXPERIMENT

In order to verify whether tactile stimulation can influence the perception of auditory spoken language without implying significant cognitive effort, we decided to run an extremely simple psychophysical experiment on the discrimination of auditory syllables /BA/ and /VA/. Tactile stimulation was presented to the thumb or finger via a specially constructed dynamic Braille display (see Apparatus, below).

We first did an auditory threshold measurement, in which subjects judged where on an auditory continuum between /BA/ and /VA/ they perceived the syllable to sound more like /BA/ or like /VA/. We also determined a tactile threshold where on a tactile continuum they felt a clear apparent movement on their thumb to be rightward versus leftward. Subjects then trained for approximately an hour by listening to exemplars of the syllables taken from the auditory continuum, coupled with the associated tactile stimulation, delivered both above and under tactile threshold. Finally, in order to assess a possible automatic influence of the tactile information on the auditory discrimination of the syllables, we ran three tests. In one test the association between auditory and tactile remained unbiased as in the training session. In two other tests, we modified the association between auditory and tactile stimuli so that it was biased as compared to the training period. We expected that the auditory discrimination threshold would be shifted by the bias induced by the tactile stimulation. In order to avoid cognitive response effects in the tests, we told subjects that the tactile stimulation could be misleading, and asked them to ignore it.

Methods

Subjects

We tested 10 normal hearing subjects (8 males, 2 females, mean age 33.2). Subjects were students, Ph.D. students and post-hoc researchers. Subjects belonged to different cultures and had different mother tongues.

Apparatus

The tactile stimulation device consisted in two side-by-side dynamic braille stimulators (B11, Metec AG, Stuttgart, Germany) each with 2 x 4 pins, creating a 4x4 array. The device was held in the hand with the Subject's thumb lying on the array (see Figure 1). The device was interfaced to a computer and controlled by a python program. The auditory stimuli were delivered through headphones at the same comfortable volume for all subjects.

Fig.1 Tactile stimulation device

Stimuli

For the auditory stimuli, we created a perceptual continuum of 13 stimuli going from the sound /VA/ to the sound /BA/. We recorded a natural male voice saying /VA/ and /BA/. We synchronized the two stimuli according to their starting points (there was silence before and after each stimulus). We spliced together segments from /VA/ and /BA/ of correspondingly shorter and longer durations so as to obtain 13 composite stimuli that sounded like /VA/ or /BA/ or some in-between composite. The splicing was done as follows. We cut the recorded stimuli /VA/ and /BA/ into four segments, containing the sounds /v/, /a/ (from the /VA/), /b/ and /a/ (from the /BA/) respectively. The file containing the sound /v/ and /b/ have the same duration. In the creation of the stimuli we only used the sound /a/ extracted from the recording /VA/, since when attached to the sound /b/ the resulting audio file seemed to be a quite natural pronunciation of the sound /BA/. In other words, we did not use the sound /a/ from the /BA/ at all in the creation of the stimuli. The 13 stimuli were then created by attaching a percentage p from the beginning of the recording of /v/ with a percentage (1-p) from the end of the sound /b/ and then attaching to this to the sound /a/ from /VA/. The percentage p varied in the 13 stimuli going from 0 to 1. We then applied a lowpass filter with a cutoff at 1850 Hz to each exemplar. We did this because based on pilot subjects' reports we had found that otherwise the auditory boundary on the continuum between perceived /VA/ and /BA/ was too sharp. Note that the low pass filtering approximates the deficit experienced by hearing impaired people.

For the tactile stimulation, we constructed 13 stimuli, in which the top line of pins stroked the thumb by moving at different speeds either from left to right or from right to left. The stimuli were constructed in the following way. Each pin stayed up for 20ms, and then it went down. To create the impression of stroking, we varied the gap in time between the moment when successive pins went up. For example for a time of 0 ms, the pins went up all together and there was no apparent stroking motion. For a time of +40 ms each successive pin came up 40 ms after the beginning of the moment when previous pin came up, giving an apparent motion from left to right. For a time of -40 ms, the motion went in the opposite direction. We used times of -40ms, -30ms, -20ms, -15ms, -10ms, -5ms, 0ms, 5ms, 10ms, 15ms, 20ms, 30ms, 40ms.

In the conditions where auditory and tactile stimuli were associated, we used all of the 13 possible elements in the auditory continuum, but these were associated with only two exemplars taken from the tactile continuum, which changed for each training session and for the last two blocks (see Procedure).

Procedure

Each Subject ran six parts: measurement of the auditory threshold, measurement of the tactile threshold, 9 blocks of auditory-tactile training, auditory-tactile test, auditory-tactile test with bias with tactile stimulation below threshold ("weak"), auditory-tactile test with bias with tactile stimulation above threshold ("strong"), measurement of the auditory threshold "post-test". After each block we asked participants to report what they perceptually heard during the block. Additionally, they could inform the experimenter at any time if they heard something that changed in their auditory perception.

The measurement of the auditory threshold included a familiarization block and a main block. In the familiarization block, the Subjects heard three stimuli (clear /VA/, clear /BA/ and the middle exemplar half way between the extremes), each one repeated three times (in random order) and they had to answer whether they heard /BA/ or /VA/. Subjects pressed the button '1' when they heard /BA/ and the button '2' when they heard /VA/. In the main block, we presented all 13 auditory stimuli in random order (each of them was presented 15 times), and Subjects responded using the same buttons to indicate whether they heard /BA/or /VA/. In this phase, we obtained a psychophysical function for each Subject, and determined his or her threshold value [μ] (the stimulus that the subject perceived 50% of the time as /BA/ or /VA/) and the measure of the slope or deviation [σ] from the threshold such that the subjects clearly perceived /BA/ or /VA/ 86% of the time. When the value of σ is smaller, the curve is steeper, which means that the discrimination of the two syllables was easier.

The measurement of the tactile threshold included two familiarization blocks and a main block. We asked the subject to press the button '1' when the subject felt that the pins moved to the left and to press the button '2' when the subject felt that the pins moved to the right. In the first familiarization block we presented the pins moving to the left with different speeds (-40ms, -30ms, -20ms, .-15ms, -10ms, -5ms), each speed repeated twice (in random order), while in the second familiarization block, we presented the pins moving to the right with different speeds (5ms, 10ms, 15ms, 20ms, 30ms, 40ms) each speed repeated twice (in random order). In the main block of measurement of tactile threshold we presented all 13 tactile stimuli in random order (each stimulus was presented 15 times), and Subjects indicated in the same way as before whether they felt the motion going from left to right or opposite. We plotted psychophysical functions for each Subject, and determined their threshold [μ] and slope [σ].

After the measurement of the auditory-only and tactile-only thresholds, the Subjects performed the audio-tactile part of the experiment (auditory and tactile stimuli were presented simultaneously) which included: one familiarization block, 9 main blocks of auditory-tactile training, one block of test, and one biased block with tactile stimulation below threshold ("weak") (see below) and one biased block with tactile stimulation above threshold ("strong") (see below). Subjects had a break after each block. For all these conditions we obtained for each subject the threshold [μ] and slope [σ].

In the familiarization block, we used the same three auditory stimuli that we used for the measurement of the auditory threshold, each one repeated three times. The extreme cases of /BA/ and /VA/ were each associated with two clearly distinguishable tactile stimuli, corresponding to the tactile stimuli that were at +3σ or -3σ from the threshold tactile stimulus.

For the 9 main blocks of auditory-tactile training Subjects were presented with all the 13 auditory stimuli; however we changed the way the two tactile stimuli were associated with them across the blocks -- first and second block: 3σ of the tactile threshold, third block: 2σ of the tactile threshold, fourth block: 1.5σ of the tactile threshold, fifth, sixth, seventh blocks: 1σ of the tactile threshold, eighth and ninth blocks: 0.5σ of the tactile threshold (see Figure 2).

The auditory-tactile test block was identical to the last session of the audio-tactile training (tactile stimuli: 0.5 σ of the tactile threshold), except for the fact that Subjects were told that tactile stimulation could be misleading and that they should ignore it.

In the biased auditory-tactile test with "weak" tactile stimulation (i.e. tactile stimulation under tactile threshold), instead of linking all the exemplars going from /BA/ to /VA/ to the tactile stimuli in a symmetrical way, we shifted the linking towards /VA/ by three points in the continuum (see Figure 3). The tactile stimulation used corresponded to the stimulations at -.5 σ and +.5 σ from the tactile threshold. Note again that we asked the Subjects to ignore the tactile inputs.

In the biased auditory-tactile test with "strong" tactile stimulation (i.e. tactile stimulation above tactile threshold) we linked auditory stimuli to tactile stimuli with the same amount of bias as in the biased with weak tactile stimulation (see Figure 3). However the tactile stimulation used was above threshold (-2.0 σ and +2.0 σ from the tactile threshold). Note again that we asked the Subjects to ignore the tactile inputs.

In a last phase of the experiment we measured the auditory threshold again without tactile stimulation (auditory threshold "post-test") as a control condition to check that any observed change in audio-tactile thresholds was really due to the tactile influence rather than some kind of auditory adaptation.

Fig. 2. Audio-tactile training. For the 9 main blocks of auditory-tactile training we used randomly all the 13 auditory stimuli, while the tactile inputs changed across the blocks: first and second block: 3σ of the tactile threshold, third block: 2σ of the tactile threshold, fourth block: 1.5σ of the tactile threshold, fifth, six, seventh blocks: 1σ of the tactile threshold, eighth and ninth blocks: .5σ of the tactile threshold.

Fig. 3. The three conditions of the test. The auditory-tactile test block was identical to the last session of the audio-tactile training (tactile stimuli 0.5 σ), except for the fact that Subjects were asked to ignore the tactile stimuli. In the auditory-tactile test with weak bias we also asked to the Subjects to ignore the tactile inputs, but this time we shifted the boundary (the point where subjects hear 50% of the time /BA/ and 50% of the times /VA/) by 3 positions, and we presented the two tactile stimuli below threshold (0.5 σ). In the auditory-tactile test with strong bias we again asked the Subjects to ignore the tactile inputs, we also shifted the boundary (the point where subjects hear 50% of the time /BA/ and 50% of the times /VA/) by 3 positions, but this time we presented the two tactile stimuli above threshold (2 σ).

Data analysis

We measured the number of times a participant responded /BA/ or /VA/ to each stimulation in each block. The frequencies of responses were fitted with cumulative normal distributions using the PsychoPy toolbox. The mean and the variance of the corresponding normal distribution were used as bias and threshold for a given block.

Paired t-tests were used to test the differences between different experimental conditions. The analysis was performed in R using the "t.test" function.

Results

Thresholds

Figure 6 shows a plot of the /BA/-/VA/ thresholds for each subject in the pre-, post- and experimental conditions, and the means of these thresholds over all subjects.

We were expecting that the thresholds for the biased conditions with 'weak' and 'strong' tactile stimulation would both be significantly higher than in the non-biased conditions (i.e. the auditory pre- and post- tests, and the test condition with no bias).

However this was not evident from the statistics. We did find a statistically significant difference between the strong bias condition and the auditory pre- and post-tests (p=0.01; p=0.005). However the difference was not significant as compared to the no-bias test condition (p=0.14). As concerns the weak bias condition, this was not significantly different from the other conditions.

Given these results, the tactile stimulation seems not to strongly influence the auditory thresholds, except when it is presented above threshold.

Fig. 6. Threshold values along the /BA/-/VA/ auditory continuum for each Subject in each condition, and overall means.

Self reports

We were expecting that in the training phase and the unbiased tests, Subjects would hear the stimuli more clearly thanks to the tactile stimulation. However this was not evident from the subjective reports. On the other hand after several blocks of training, or during the test conditions, some Subjects heard sounds intermediate between /BA/ and /VA/, like for example /vga/ or /nga/ (e.g. Subject 4). Additionally almost all subjects reported hearing sounds that were completely different from the stimulus sounds. For example, some Subjects heard the sound "clack" (e.g. Subject 8) or 'eco' (e.g. Subject 5). See the Conclusion for more discussion of this finding.

Discussion

This experiment was designed to test if it is possible to obtain truly perceptual facilitation of auditory speech information by the use of a learned auditory-tactile association. Unfortunately our results are not clearcut. First, any tactile effect we observed was rather weak, and second it only occurred for above-threshold tactile bias. We cannot exclude the possibility that Subjects disobeyed our instructions to not focus on the tactile information. Thus, we cannot exclude the possibility that subjects made use of the tactile information in a cognitive fashion, instead of integrating it perceptually with the auditory signal (for a similar distinction for auditory and visual integration see the 'bimodal speech perception' theory of Massaro, 1987; Massaro and Cohen, 1983).

Concerning the finding that most of our Subjects heard peculiar sounds which were neither /VA/ nor /BA/: We observed that this effect occurred also in the purely auditory post-test, and so we hypothesized that it could simply be due to selective speech adaptation (Eimas and Corbit, 1973), a purely auditory phenomenon known to occur for repeated presentation of syllables. We further confirmed in an independent pilot test that a Subject who ran the experiment with no tactile stimulation at all also experienced such auditory deformations.

It might be argued that the reason we did not obtain interference from the below-threshold bias and only a small interference in the above-threshold bias was that Subjects were not sufficiently trained: after all the association between our auditory and tactile stimuli was completely arbitrary, and it is known that in such cases extensive training is necessary to find an interaction.

However it should be noted that even if, with further training, we had obtained stronger effects, and in particular an effect of the sub-threshold tactile bias, we could still not be fully confident that the phenomenon was purely perceptual. Even sub-threshold tactile information might at times during the experiment reach the subject's awareness and influence his or her decisions in a cognitive rather than perceptual way. For this reason we decided to use a tactile analogy of the McGurk effect (Experiment 2), which provides a perceptually obvious instance of multimodal integration. To avoid the problem of auditory adaptation due to repetition of the same syllables, we decided to increase the number of syllables we used.

EXPERIMENT 2: AUDIO-TACTILE MCGURK

After the absence of the expected result in Experiment 1, and given the difficulty of using that technique for assessing if there can be truly perceptual auditory-tactile integration, we decided to use a tactile analog of the McGurk effect to attempt to obtain a subjectively obvious perceptual effect demonstrating integration of tactile and auditory signals.

In the original McGurk effect, a visual /GA/ accompanied by an acoustic /BA/ is often perceived as /DA/; a visual /BA/ accompanied by an acoustic /GA/ is sometimes perceived as /BGA/ (McGurk & MacDonald, 1976). If we could find evidence for an audio-tactile McGurk type illusion, where a new syllable is perceived, then this would argue for true audio-tactile integration rather than unimodal selection of either the acoustic or the tactile stimulus (see 'bimodal speech perception' theory in auditory-visual domain, Massaro, 1987; Massaro and Cohen, 1983).

As far as we know, two previous studies have tried to demonstrate auditory-tactile integration using a paradigm close to the McGurk effect in the audio-tactile domain, pairing repeated auditory syllable to tactile stimulation (Fowler and Dekle, 1991; Sato et al. 2010). Both studies did not report a perceptual experience equivalent to the original McGurk effect (perception of a different syllable from the four presented). Fowler and Dekle (1991) found an equivalent of the McGurk effect in only in one out of 7 Subjects, even if there was evidence for a strong influence of the tactile stimulation on auditory perception of syllables (and vice versa). Sato et al. (2010) found no clear evidence for McGurk type illusion, since they found perception of the same syllable in both unimodal and bimodal condition. However, these two studies differ from ours for two reasons. First of all, they did not use an arbitrary code, as tactile aids do, but a method called Tadoma (Alcorn, 1932) where, in its original version, a deafblind person places their thumb lightly on the lips of the talker while the other fingers fan out over the face and neck. Secondly, in both studies there was no training phase. We hoped that, even though we used an arbitrary tactile stimulation code that Subjects had no prior familiarity with, after extensive training there might be evidence for auditory-tactile integration and that this integration would take place in a way similar to what happens in the original McGurk effect.

We therefore extensively trained two Subjects to associate four syllables with four corresponding tactile patterns. In a subsequent test, we presented the Subjects with the learnt auditory syllables, except now they were accompanied either with the previously associated tactile patterns (congruent trials), or with tactile patterns that were not associated with them (incongruent trials). We expected that in a way analogous to the McGurk effect, Subjects in the incongruent trials would perceive a completely "new" syllable or a syllable "inbetween" the auditory and tactile syllables.

Methods

Subjects

2 male subjects (mean age 34.5 yrs). Russian and Italian native speakers (experimenters).

Stimuli and Apparatus.

Auditory stimuli: We chose four syllables (/BA/, /VA/, /GA/, /DA/) that were easily confused, as measured by a classic confusion matrix for phonemes (Miller and Nicely, 1955). We used 10 different instances of each syllable (40 stimuli overall). The syllables were presented at 4 different levels of volume ranging from just audible to very clearly audible. In the blocks of training without tactile stimulation the level of the volume was provided at the third, intermediate loudness level.

Tactile stimuli: we used the same device as in Study 1. Tactile stimulation was always above threshold (time delay 60 ms, and 60 ms pin-up time, see Study 1), so its direction could be easily judged by the subjects. To make four distinguishable tactile stimuli (up, down, left and right), we used the external pins of the 4x4 array (see Fig. 8).

We trained the two Subjects to associate each syllable to one tactile stroking direction. The Subjects used two different sets of keys on the QWERTY keyboard for responding: Subject 1 used keys corresponding to the position on the thumb that was being stimulated: W (for /DA/), A (for /BA/), S (for

/GA/), D (for /VA/); whereas Subject 2 used keys corresponding to the apparent motion of the pins: W (for /VA/), A (for /DA/), S (for /BA/), D (for /GA/). (see Fig. 8 and 9)

Fig. 8 Stroking directions of the pins.

W

A S D

Fig. 9 Keyboard button layout.

Procedure

Each subject did daily training and a test at the end of the day. For the training sessions (divided into 'active' and 'passive' sessions), the Subject sat in front of the computer wearing headphones with his thumb resting on the tactile device. Subjects had a break after each session of 200 trials of training.

In the 'passive' training sessions, when the Subject pressed a key on the keyboard, the computer randomly delivered one of the four sounds, and the Subject responded using the allocated keys on the keyboard (see Stimuli). All 10 instances of each of the 4 syllables were presented. The passive training included 4 condition of the training: auditory only, tactile only, auditory plus tactile, no answer (Subjects mentally registered the syllable they heard, but they did not press a button). Four out of five times the auditory presentation was also accompanied by corresponding tactile stimulation: in other words, one fifth of the time, there was no tactile stimulation. This was done so as to force subjects not to rely only on the (unambiguous) tactile stimulation, and to pay attention also to the auditory stimulation.

In the 'active training' the participant read the syllable displayed on the screen and at the same time felt the tactile stimulation (about 1000 trials per day).

The test included three sequences of 10 trials which were congruent (e.g. /BA/ presented 10 times with the learned tactile stimulation) and three sequences of trials which were incongruent (each syllable associated with each of the possible incongruent tactile stimulations), in a random order. There were thus 30 congruent and 30 incongruent trials for each auditory/tactile combination, in random order. In this way we had three times more congruent stimuli in the test than incongruent stimuli. We maintained this imbalance between the congruent and incongruent trials in order to avoid invalidating participants' learnt associations. Subjects pressed a key on the keyboard to obtain each successive stimulus.

To test for the existence of a perceptual illusion, the Subject listened to 10 successive examples of a given auditory/tactile combination, and determined what sound he thought he heard. In order to check for the existence of a perceptually obvious McGurk-like phenomenon, before the end of the sequence of 10 presentations, the Subject removed his thumb from the device and looked at the tactile stimulation presented in order to check if the tactile stimulation that he saw corresponded to the auditory percept that he had had in the previous trials.

Overall, Subject 1 did 6 days of training, 4-5 hours per day (around 50.000 trials), while Subject 2 did 6 days of training, 1.5-2 hours per day (around 15.000 trials).

Each Subject also ran one session of test per day (240 trials) randomly mixed with half consisting of congruent trials (tactile stimulation corresponds to the trained auditory syllables), and half of incongruent trials (tactile stimulation does not correspond to the trained auditory syllable). In the incongruent trials, all 12 incongruent combinations of auditory and tactile stimulation were presented.

Results and Discussion

Our expectation was that for the incongruent trials Subjects would perceive a 'new' syllable or a totally different syllable from the presented /BA/, /VA/, /DA/ or /GA/. However, nothing changed in Subjects' perception of the syllables either in the congruent or incongruent trials.

In this qualitative study we wanted to find a proof of an audio-tactile integration at the syllable level. We hoped to find an equivalent of the McGurk effect, which we consider to be a good example of truly perceptual multisensorial integration. Our Subjects did not report any perceptual change in the heard syllables when the tactile stimulation was not congruent with the learnt one. We therefore again have no proof for audio-tactile integration. Even after the long period of training, we found that tactile input from our device did not influence the auditory perception of syllables.

Conclusion

In the introduction we pointed out that one problem that may explain why hearing impaired users have never adopted audio-tactile speech prostheses may be that they do not provide an impression of "hearing better". Indeed, even though previous studies on multisensory integration have shown that auditory-tactile integration is common in the non-speech domain, there is no clear proof in the literature that audio-tactile integration occurs in speech. Here our purpose was therefore to attempt to gather evidence that such integration was after all possible in the speech domain.

We used apparatus which was improved as compared to devices that had been used in the past. In particular we avoided the problem of unpleasantness and habituation of the skin that is produced by vibratory or electric stimulation: we did this by using an apparent tactile motion stimulus. Similarly, the problems of cognitive effort and adequacy of the information provided by the tactile signal were obviated

by using an extremely simple situation involving only two syllables (4 in Expt 2), each associated with one tactile pattern.

We ran two studies. Study 1 was a very simple psychophysical procedure to check whether the perception of auditory syllables in an auditory continuum (from a clear /BA/ to a clear /VA/) would be influenced by simultaneous presence of a previously learnt compatible or incompatible tactile stimulation presented above and below threshold. We found that tactile stimulation modified the threshold of auditory discrimination of the syllables /BA/ and /VA/ only when it was above threshold. Unfortunately therefore we cannot exclude that subjects used a strategy of selecting the most reliable unimodal information source, instead of experiencing true perceptual integration between the auditory and tactile modalities. We observed that many participants heard curious sounds which might have been evidence of perceptual modification of the acoustic stimulation by the tactile stimulation. However a pilot experiment with only auditory input suggested that the effect was actually simply due to selective speech adaptation (Eimas and Corbit, 1973) rather than a consequence of the tactile stimulation. For this reason we decided to run a second qualitative study (Study 2) where we used a tactile analog of the McGurk effect to attempt to find proof of the existence of a perceptual effect of the tactile stimuli on the auditory. However, after extensive training with a particular auditory-tactile association, we found that Subjects' auditory syllable perception did not change when the tactile stimulation was incongruent as relative to the learned association. In other words, we did not find an audio-tactile equivalent of the audio-visual McGurk effect. Both our studies thus failed to find any evidence for the existence of truly perceptual auditory-tactile integration at the syllable level.

In conclusion, although the literature shows that audio-tactile integration is a common phenomenon in the non-speech domain, our experiments did not provide any evidence of such multimodal audio-tactile integration in the speech domain.

Further studies are required in order to check if this audio-tactile integration occurs in the speech domain at a perceptual level. If such integration can be found this would considerably favour the development of tactile devices that can help hearing impaired people to better understand speech and to have the experience of "hearing better".

Acknowledgments

This work is funded by the ERC Grant FEEL Number 323674.

We wish to thank Ramakanth Singal for his fundamental contribution in the construction of the device (See Apparatus).

References

Alcantara J., Blamey P.J and Clark G. M. ***Tactile-Auditory Speech Perception by Unimodally and Bimodally Trained Normal-Hearing Subjects*** **J Am Acad Audiol 4 :98-108 (1993)**

Alcorn S. ***The Tadoma method*** **Volta Review, 34, 195-198 (1932)**

Brooks P.L., Frost B.J., Mason J.L.and Chung K. ***Acquisition of a 250-word vocabulary through a tactile vocoder*****. J. Acoust. Soc. Am. 77 (4), April 1985**

Brooks PL, Frost BJ, Mason JL, Gibson DM ***Continuing evaluation of the Queen's University tactile vocoder . I:Identification of open set words.*** **J Rehabil Res Dev 23(1) :119-128, 1986**

Caetano G and Jousmäki V. ***Evidence of vibrotactile input to human auditory cortex*** **NeuroImage Volume 29, Issue 1, 1 January 2006, Pages 15–28**

Calabrese I. and Olivetti Belardinelli M. (1997). Musical Abilities in Deaf Children: Assessment Development and Training. In: A. Gabrielsson (Ed.). Third Triennal ESCOM Conference: Proceedings. Uppsala, June

Cowan R.S.C., Blarney ..J., Galvin K. L.,Sarant J.Z., Alcantara J. I. and Clark G. M. ***Perception of sentences, words, and speech features by profoundly hearing-impaired children using a multichannel electrotactile speech processor*** **J. Acoust. Soc. Am. 88 (3), September 1990**

Darrow Alice-Ann ***Music and the Hearing Impaired: A Review of the Research with Implications for Music Educators*** **Applications of Research in Music Education Spring 1989 7: 10-12**

Eimas P. D. and Corbit J. D. ***Selective Adaptation of Linguistic Feature Detectors.*** **Cognitive Psychology 4, 99-109 (1973)**

Engelmann S. and Rosov R. J. " ***Tactual Hearing Experiment with Deaf and Hearing Subjects"*** **Research bulletin, Vol. 14, n.5 June 1974, 49 p.**

Fowler and Dekle, 1991 ***Listening with Eye and Hand: Cross-modal Contributions to Speech Perception*** **HCJ8'IUns Laboratories Status &porl 00 Speech &~arch**
1991, SR-107/108, 63-80

Foxe J.J., Multisensory integration: frequency tuning of audio-tactile integration, Current Biology Vol 19 No 9
R374, 2009

Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM. ., *Auditory-Somatosensory Multisensory Processing in Auditory Association Cortex: An fMRI Study*, J Neurophysiol 88: 540–543, 2002; 10.1152/jn.00694.2001

Galvin, K.L., Ginis J., ; Cowan R. S. C., Blamey P. J., Clark G. M. *A comparison of a new prototype Tickle Talker with a Tactaid 7* Australian and New Zealand Journal of Audiology, May, 23(1), 18-16 (2001).

Gick B. & Derrick D. *Aero-tactile integration in speech perception* Nature 462, 502-504 (26 November 2009)

Gillmeister H. and Eimer M., *Tactile enhancement of auditory detection and perceived loudness*, Brain Res., 2007 July 30;1160:58-68. Epub 2007 Mar 20. 7-12 1997, p. 87-90

Jousmäki V. and Hari R. , 1998 *Parchment-skin illusion: sound-biased touch* R190 Current Biology, Vol 8 No 6

Kassubba T., Menz M. M. , Ro B. and Siebner H. R. *Multisensory Interactions between Auditory and Haptic Object Recognition.* Cerebral Cortex May 2013;23:1097±1107

Lynch M. P ., Eilers E .R.MS., Oller K.,and Lavoie L., *Speech perception by congenitally deaf subjects using an electrocutaneous vocoder* Journal D . of Rehabilitation Research and Development Vol . 25 No . 3 Summer 1988 Pages 41 -50

Massaro, D. *Speech perception by ear and eye: A paradigm for psychological inquiry.* Hillsdale, NJ: Erlbaum (1987)

Massaro, D.W. & Cohen M.M. E*valuation and integration of visual and auditory information in speech perception* Journal of Experimental Psychology Journal of Experimental Psychology: Human Perception and Performance 1983, Vol. 9, No. 5, 753-771

McGurk & MacDonald, 1976 *Hearing lips and seeing voices* Nature, Vol 264(5588), pp. 746–748

Miller G. A. and Nicely P.E. *An Analysis of Perceptual Confusions Among Some English Consonants* J. Acoust. Soc. Am. 27, 338 (1955)

Olivetti Belardinelli M. *Looking for Effective Accessibility of Sonificated Spatial Information to Blind People: Four Experiments* FORUM ACUSTICUM 2011 27. June - 1. July, Aalborg

Ro T. , Hsu J., Yasar N.E., Elmore L.C., Beauchamp M.S. *Sound enhances touch perception,* Exp Brain Res (2009) 195:135–143

Sato M., Cavé C. and Ménard L. *Auditory-tactile speech perception in congenitally blind and sighted adults* Annie Brasseur Neuropsychologia 48 (2010) 3683–3686

Schürmann, Martin & Caetano, Gina & Jousmäki, Veikko & Hari, Riitta. *Hands help hearing: Facilitatory audiotactile interaction at low sound-intensity levels* The Journal of the Acoustical Society of America 2004. Volume 115, Issue 2. 830-832. ISSN 1520-8524

Sicilley, P. L. *Evaluation of a vibrotactile auditory prosthetic device for the profoundly deaf* unpublished Master's thesis(1980), Queen's University, Kingston, Canada

Soto-Faraco S, Spence C. and Kingstone A. *Congruency effects between auditory and tactile motion: Extending the phenomenon of cross-modal dynamic capture* Cognitive, Affective, & Behavioral Neuroscience 2004, 4 (2), 208-217

Weisenberger J. M. and Broadstone S. M.*Evaluation of two multichannel tactile aids for the hearing impaired* J.Acoust. Soc. Am., Vol.86, No.5, November 1989

Yarrow K., Haggard P. and Rothwell J.C. *Vibrotactile–Auditory Interactions are Post-Perceptual* Perception, 2008, volume 37, pages 1114 ^ 1130

Yau J. M. , Weber A. I. and Bensmaia S. J. *Separate mechanisms for audio-tactile pitch and loudness interactions, Frontiers in Psychology*| Perception Science October 2010, Volume 1, Article 160

Yau J. M., Olenczak J. B. , Dammann J. F. and Bensmaia S. J. *Temporal frequency channels are linked across audition and touch* Curr Biol . 2009 April 14; 19(7): 561–566.