

BIANCA SISINNI, BARBARA GILI FIVELA, MIRKO GRIMALDI

L'elaborazione preattentiva del tratto di durata in suoni non nativi: uno studio elettrofisiologico

Native Italian listeners do not use vowel duration to distinguish word meaning. Nonetheless, similarly to Spanish listeners (e.g., Chládková et al., 2013), they are supposed to be able to use duration to discriminate vowel pairs that belong to nonnative or second language (L2) systems, especially when similar to one native sound (Desensitization Hypothesis; Bohn, 1995). The present study investigates how Italian listeners elaborate duration in different L2 vowels, i.e., similar and new with regard to the native systems, by means of the Mismatch Negativity, an electrophysiological component that reflects a pre-attentive change detection of a deviant sound in a sequence of frequent sounds. The results show that Italian listeners can make use of the duration cue, irrespectively of the vowel sound. In addition, the use of double deviants, i.e., vowel formant frequency and duration differences, carried by a single stimulus (e.g., Sussman, Winkler, 2001) allowed us to observe a sequential processing of both duration and frequency.

Key words: Second language perception, vowel duration, Mismatch Negativity, double deviants.

Introduzione

La relazione di similarità/differenza fra il sistema nativo (L1) e un sistema non nativo (NN) o una seconda lingua (L2) è una delle cause principali delle difficoltà nel percepire suoni NN/L2 che sono caratterizzati da correlati non presenti o utilizzati in maniera differente nella L1 (Escudero, Benders & Lipski, 2009). Ad esempio, i parlanti nativi del giapponese hanno difficoltà nel discriminare i fonemi /r-l/ in quanto non sono in grado di percepire la differenza della terza formante (F3) fra le due consonanti (Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann & Siebert, 2003).

Spesso i parlanti nativi di L1, nel discriminare contrasti NN/L2, si basano su dimensioni acustiche che per i parlanti nativi di L2 sono ridondanti o secondarie. Ad esempio, vi sono sistemi linguistici in cui il tratto di durata (consonantica e/o vocalica) ha una valenza fonologica e, pertanto, vengono definiti *quantity systems*. Ad essi si oppongono i *quality systems*, nei quali la differenza di durata non determina variazioni di significato (Nenonen, Shestakova, Huotilainen & Näätänen, 2005). L'uso della durata nella discriminazione di contrasti vocalici sembra differire fra i parlanti di lingue *quality* vs. i parlanti di lingue *quantity*. Ad esempio, i parlanti nativi dell'inglese distinguono le vocali /i:/-/ɪ/ principalmente o esclusivamente sulla base delle differenze formantiche fra le due vocali mentre è stato osservato che parlanti nativi dello spagnolo, del cinese mandarino (Flege, Bohn & Jang, 1997), del portoghese (Rauber, Escudero, Bion & Baptista, 2005), del catalano (Cebrian, 2006), del polacco (Bogacka, 2004), del russo (Kondaurova, Francis, 2008) e del giapponese (Morrison, 2002) basano la discriminazione del contrasto, assi-

milato alla vocale nativa /i/, sul tratto di durata, nonostante esso non abbia una valenza fonologica nei rispettivi sistemi nativi. Questi risultati sembrerebbero confermare l'ipotesi formulata da Bohn (1995), secondo la quale i parlanti sarebbero meno o non affatto sensibili (*desensitization hypothesis*) a differenze spettrali fra due vocali NN o L2 in regioni dello spazio acustico che contengono una sola vocale nativa. Conseguentemente, baserebbero la discriminazione della coppia NN o L2 sulla differenza di durata, essendo una dimensione psico-acusticamente saliente.

La rilevanza della durata nella discriminazione di contrasti nativi o non nativi è stata osservata anche attraverso lo studio della *Mismatch Negativity* (MMN), una componente elettrofisiologica che riflette la discriminazione preattentiva di stimoli uditivi, il cui picco massimo in ampiezza si registra attorno ai 100-250ms dalla presentazione dello stimolo. La MMN si elicitava, fra gli altri, nel classico paradigma *oddball* in cui una serie di stimoli frequenti, detti standard, viene interrotta da stimoli meno frequenti, detti devianti, che differiscono dallo standard per una (o più) dimensione, come le frequenze formantiche o la durata.

Ylinen, Houtilainen & Näätänen (2005), in uno studio sull'elaborazione della durata consonantica in parlanti nativi del finlandese (*quantity system*), hanno ipotizzato l'esistenza di meccanismi neurali indipendenti e paralleli per la qualità e la quantità (durata) consonantica. Infatti, in risposta ad uno stimolo deviante che differisce dallo standard sia per qualità che per durata, gli autori hanno, riscontrato la presenza di una MMN unica che è risultata pari alla somma delle MMN generate dai devianti singoli per qualità o per durata (MMN additiva). In questo modo, hanno confermato la presenza di processi neurali indipendenti che si sommano fra loro per l'elaborazione del doppio deviante (si veda successivamente la definizione di doppio deviante).

Per ciò che concerne i suoni non nativi, Lipski, Escudero & Benders (2012) hanno confrontato le MMN in parlanti nativi dello spagnolo che discriminano il contrasto olandese /a/-/ɑ/, caratterizzato da differenze spettrali, e il contrasto /a:/-/a/, caratterizzato dalla sola differenza di durata (dove /a/ è un fonema nativo). Le MMN osservate per il contrasto spettrale hanno un'ampiezza minore rispetto al contrasto di durata e gli autori hanno quindi dedotto che l'elaborazione del tratto di durata nella discriminazione di vocali simili avviene anche a livello preattentivo.

La stessa elaborazione preattentiva sembra essere modulata dal rapporto di similarità/dissimilarità fra i suoni L2 e il sistema nativo. Sia Nenonen et al. (2005) che Chládková, Escudero & Lipski (2013) hanno osservato che le MMN relative all'elaborazione della durata sono maggiori per stimoli vocalici che differiscono dal sistema nativo, ovvero non sono categorizzabili rispetto ad esso e possono essere considerati *nuovi*, che per stimoli *simili*, riconducibili invece a categorie native. In entrambi gli studi viene ipotizzato che in parlanti di un *quality system* il sistema nativo inibisce l'elaborazione del tratto di durata per quei suoni che possono essere categorizzati rispetto al sistema nativo stesso, mentre non esercita alcuna azione inibitoria per i suoni non categorizzabili. In sostanza, poiché nel sistema nativo il tratto di durata vocalica non è presente, esso non avrebbe alcun effetto durante l'elaborazione di vocali categorizzabili come native. Al contrario, la durata sarebbe facilmente elaborata per quei suoni che non sono categorizzati in maniera

“coerente” dal sistema nativo. Questo accadrebbe sia per parlanti con esperienza in L2 (Nenonen et al., 2005) che per parlanti monolingue, senza alcuna conoscenza della L2 (Chládková et al., 2013).

In entrambi gli studi citati, lo stimolo deviante differisce dallo standard per una sola dimensione, ovvero il tratto di durata. Nenonen et al. (2005) hanno confrontato le MMN di parlanti russi in risposta al contrasto finlandese simile (categorizzabile con il sistema nativo) /ka/-/ka:/ con quelle in risposta al contrasto finlandese *nuovo* /kæ/-/kæ:/, mentre Chládková et al. (2013) hanno comparato le risposte MMN ottenute, fra gli altri, in parlanti spagnoli che percepivano i contrasti estoni *simile* /æ/-/æ:/ vs. *nuovo* /ɤ/-/ɤ:/.

Diversamente dagli studi di Nenonen et al. (2005) e Chládková et al. (2013), Ylinen, Uther, Latvala, Vepsäläinen, Iverson, Akahane-Yamada & Näätänen (2010) utilizzano stimoli standard e devianti che differiscono fra loro non solo per la durata ma anche per valori formantici, ovvero le vocali inglesi /i:/-/ɪ/ (nelle parole /b_t/). L'obiettivo dello studio è osservare le risposte MMN in parlanti nativi del finlandese per verificare se un training percettivo possa modificare le modalità di elaborazione del tratto di durata. Oltre alla presenza di stimoli che differiscono fra loro in due aspetti, gli autori hanno previsto l'utilizzo invertito di standard e deviante, per cui in una condizione lo standard è la vocale L2 /i:/ e nell'altra la vocale L2 /ɪ/. Anche in questo caso, il grado di similarità/dissimilarità rispetto alle categorie native sembra giocare un ruolo determinante. Essendo la vocale L2 /i:/ simile alla vocale nativa /i/, le MMN elicitate quando la vocale L2 /i:/ è utilizzata come standard sono di ampiezza maggiore rispetto alle MMN elicitate quando lo standard è la vocale L2 /ɪ/. Il contesto di elicitazione, quindi, sembra avere un ruolo cruciale in quanto utilizzare stimoli standard simili alle categorie native facilita l'attivazione delle rispettive rappresentazioni presenti nella memoria a lungo termine necessarie per una più regolare elicitazione della MMN (si veda anche Näätänen, 2001).

Pertanto, utilizzare come standard uno stimolo riconducibile ad una categoria nativa sembra essere la condizione ideale per un'elicitazione più corretta della risposta MMN. Inoltre, tale utilizzo sembra rispecchiare meglio ciò che accade nella realtà, in quanto stimoli percettivi in entrata vengono comparati con le categorie native dei parlanti (e non con stimoli nuovi o non presenti).

Tuttavia, come accennato in precedenza, ciò che sembra essere rilevante nel lavoro di Ylinen et al. (2010) è il ricorso a “doppi devianti”, *double deviants*, ovvero devianti che differiscono dallo standard non solo per durata ma anche per valori formantici.

Più specificamente, i *double deviants* (Sussman, Winkler, Ritter, Alho & Näätänen, 1999; Sussman, Winkler, 2001; Jaramillo, Ilvonen, Kujala, Alku, Tervaniemi & Alho, 2001; Wang, Datta & Sussman, 2005; Océák, Winkler & Sussman, 2008) possono definirsi come due deviazioni i cui onset occorrono in un intervallo temporale di circa 200 millisecondi (ms), ovvero la finestra temporale di integrazione (*Temporal Window of Integration*, TWI)¹. Sussman, Winkler (2001) affermano che, indipendentemente dal fatto che le deviazioni siano veicolate da un unico stimolo o da due stimoli differenti, i

¹ La TWI può variare fra i 150ms (Wang et al., 2005) e i 200-250ms (Horváth, Czigler, Winkler & Teder-Sälejärvi, 2007).

doppi devianti elicitano una sola MMN in quanto le due deviazioni vengono elaborate parallelamente ed, infine, integrate, come fossero un unico evento deviante. Tuttavia, quando i doppi devianti sono presentati in sequenze di stimoli nelle quali compaiono anche “devianti singoli”, ovvero stimoli che differiscono dallo standard per una delle due deviazioni dei doppi devianti, allora saranno elicitate due MMN successive, in quanto la seconda deviazione coinciderebbe con un’informazione nuova rispetto al contesto (ovvero la presenza di devianti singoli) e, per questo, rilevante. La presenza/assenza di devianti singoli, pertanto, determinerebbe il modo in cui i doppi devianti vengono elaborati, se come un evento singolo o come due eventi successivi. La presenza di due MMN elicitate da due deviazioni veicolate da un unico stimolo in presenza di devianti singoli è stata riscontrata anche nel lavoro di Oeck et al. (2008), nella condizione che egli definisce *Isochronous Combined Deviation*. Al contrario, Czigler, Winkler (1996) sostengono che, indipendentemente dalla presenza di devianti singoli, se le deviazioni sono veicolate da uno stimolo unico, una sola MMN sarà elicitata.

Recentemente, Althen, Huotilainen, Grimm & Escera (2016) hanno verificato se le MMN elicitate in paradigmi *multi-feature* (Näätänen, Pakarinen, Rinne & Takegata, 2004)² differissero da quelle elicitate in classici paradigmi *oddball*, utilizzando devianti singoli in frequenza e intensità e doppi devianti in “frequenza+intensità”. I risultati hanno dimostrato che le ampiezze e le latenze delle MMN ottenute nel paradigma *multi-feature* non differiscono da quelle elicitate nel paradigma *oddball*, caratterizzate in entrambe le condizioni dalla presenza di due picchi. Inoltre, hanno osservato la presenza di due MMN successive, a differenza di quanto osservato in altri studi in cui il doppio deviante “frequenza+durata”, ha elicitato una sola MMN (Wolff, Schröger, 2001; Paavilainen, Mikkonen, Kilpeläinen, Lehtinen, Saarela & Tapola, 2003).

L’obiettivo del presente lavoro è quello di indagare se il tratto di durata nella discriminazione di vocali viene elaborato preattentivamente da parlanti italiani (varietà di italiano salentino, IS). L’italiano è un sistema in cui il tratto di durata ha valenza fonologica per le consonanti ma non per le vocali. L’italiano è, quindi, una lingua *quantity*, in quanto comprende in sé il tratto di durata, ma non per le vocali. Un primo obiettivo di questo lavoro è verificare se il tratto fonologico di lunghezza porta i parlanti italiani ad elaborare le vocali lunghe come i parlanti nativi di lingue *quantity* “pure”, come il finlandese (Ylinen, Houtilainen & Näätänen, 2005), per i quali è stata osservata la presenza di meccanismi neurali indipendenti per durata e qualità formantica, oppure se esso non ha effetto sull’elaborazione delle vocali, per le quali conseguentemente ci si aspetterà un’elaborazione unica e sequenziale delle due caratteristiche dello stimolo. Tale obiettivo è in linea con il ricorso ai doppi devianti, come si può leggere successivamente.

In linea con Nenonen et al. (2005), Lipski et al. (2012) e Chládková et al. (2013), ipotizziamo che il tratto di durata venga elaborato in funzione della relazione inter-linguistica fra gli stimoli devianti e lo stimolo standard. Lo stimolo standard corrisponde alla vocale nativa /a/ (utilizzata come standard per una migliore condizione di elicitazione).

² Nel paradigma *multi-feature*, i 5 devianti differiscono dallo standard per una sola caratteristica (ad es., intensità, durata, ecc.) e vengono presentati alternatamente ad esso così che ogni deviante possa rafforzare la rappresentazione dello standard rispetto al deviante successivo.

zione della MMN, Ylinen et al., 2010) e gli stimoli devianti sono le vocali dell'inglese americano (AE) /ɑ/ e /ʌ/. La vocale nativa /o/ è utilizzata come controllo. Studi precedenti (Escudero, Sisinni & Grimaldi, 2014; Sisinni, Escudero & Grimaldi, 2013) hanno osservato che la vocale L2 deviante /ɑ/ viene percepita da ascoltatori italiani come molto simile alla vocale nativa /a/, mentre la vocale L2 /ʌ/ viene percepita come differente, nuova, rispetto al sistema nativo (non categorizzata in maniera univoca con nessun fonema nativo, Best, Tyler, 2007; Elvin, Escudero, 2014). Pertanto, ipotizziamo quanto segue. Relativamente alle sole differenze formantiche fra le vocali devianti e la vocale standard, la vocale L2 simile /ɑ/ elicerà MMN poco ampie (difficoltà di discriminazione rispetto alla vocale L1 /a/), la vocale L2 nuova /ʌ/ elicerà MMN elevate (suono nuovo facilmente discriminabile rispetto al sistema nativo), la vocale L1 /o/ elicerà le MMN maggiori (vocale nativa discriminabile rispetto alla vocale standard). Relativamente all'elaborazione della durata, invece, ipotizziamo che la vocale L2 /ɑ/ elicerà MMN poco ampie, in quanto il sistema nativo inibirà l'elaborazione del tratto di durata per la vocale L2 simile, mentre l'inibizione non avrà effetto sulla vocale L2 nuova /ʌ/, per la quale le ampiezze della MMN saranno quindi maggiori (Nenonen et al., 2005; Chládková et al., 2013). Infine, non ci aspettiamo l'elaborazione del tratto di durata per la vocale nativa deviante /o/ (controllo), in quanto le differenze formantiche fra le due vocali dovrebbero essere sufficienti per la discriminazione del contrasto e l'elaborazione della durata dovrebbe essere inibita dal sistema nativo.

Inoltre, in una versione modificata del paradigma *multi-feature* (Althen et al., 2016), si utilizzeranno doppi devianti, in cui la distanza temporale fra gli onsets delle due deviazioni è compresa nella TWI (< 200ms). Il ricorso ai doppi devianti (L2 /ɑ/-/ʌ/ e L1 /o/ con durate intermedie e lunghe rispetto alla durata dello stimolo standard L1 /a/, si veda 2.2. per i dettagli) e ai devianti singoli (L2 /ɑ/ e /ʌ/ e L1 /o/ con durata pari a quello dello standard) ci permetterà di verificare se le due deviazioni di "frequenze formantiche + durata" siano elaborate come un evento unico (una MMN), frutto dell'integrazione di due processi cognitivi indipendenti e paralleli (Ylinen et al., 2005), o come due eventi successivi (due MMN).

2. Metodologia

2.1 Soggetti

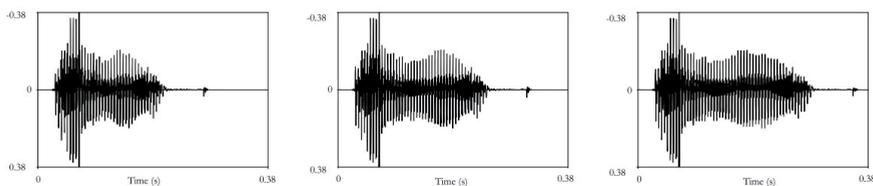
Quattordici soggetti hanno preso parte allo studio (8 uomini, età media 22,6; s.d. 1,6; 2 mancini) dopo aver preso visione e aver sottoscritto il modulo di consenso informato. Tutti i soggetti sono parlanti monolingue di italiano provenienti dal Salento, nati e cresciuti da genitori salentini. Ai soggetti è stato chiesto di indicare il loro livello di conoscenza di altre lingue, diverse dall'italiano, su una scala da 0 (nessuna conoscenza) a 7 (conoscenza nativa). In media, indipendentemente dalla lingua, hanno indicato il valore 4. Un soggetto è stato eliminato dall'analisi dei dati per problemi tecnici. Lo studio è in linea con la Dichiarazione di Helsinki (2008) ed è stato approvato dal comitato etico dell'Università del Salento.

2.2 Stimoli

Gli stimoli sono costituiti da pseudo-parole semi-sintetiche la cui struttura è /bVb/, dove V rappresenta le vocali dell'AE /ɑ/ e /ʌ/ e le vocali dell'italiano /a/ e /o/. La pseudo-parola originale è stata prodotta da una parlante dell'AE, ma le frequenze formantiche delle vocali sono state manipolate per ottenere le vocali target dell'AE e dell'IS, sulla base di valori formantici reali.

La durata della vocale standard è pari a 180ms e la durata totale dello stimolo è pari a 303ms. La durata delle vocali degli stimoli devianti, L2 /ɑ/, /ʌ/ e L1 /o/, è stata modificata allungando la parte stabile delle vocali, al di fuori delle transizioni, in modo da garantire che l'unica differenza percepibile sia quella della durata della vocale (Figura 1). Per ogni stimolo deviante, tre durate sono state ottenute: 180ms (standard), 218ms (intermedia), 255ms (lunga)³. I devianti la cui durata è pari a quella dello stimolo standard sono considerati come devianti "singoli" mentre i devianti che hanno durata intermedia e lunga come *double deviants*. La distanza temporale degli onsets delle deviazioni nei doppi devianti è minore della TWI (180ms).

Figura 1 - Da sinistra, la vocale L2 simile /ɑ/ con durata standard (180ms), intermedia (218ms) e lunga (255ms)



La frequenza fondamentale è stata tenuta costante fra gli stimoli (202 Hz) così come l'intensità (70 dB).

I soggetti, ai quali non è stata specificata la lingua di appartenenza degli stimoli, non erano consapevoli che alcuni stimoli appartenessero all'inglese e altri all'italiano.

2.3 Procedura sperimentale

È stata utilizzata una versione modificata del paradigma *multi-feature optimum 3* (Nataanen et al., 2004). Lo stimolo contenente la vocale L1 /a/ di durata pari al 180ms è lo stimolo standard mentre gli altri 9 stimoli sono i devianti che differiscono dallo standard per i soli valori formantici (devianti singoli) o per formanti e durata (*double deviants*). Gli stimoli sono stati presentati attraverso casse audio posizionate ai lati della postazione del soggetto ad un volume accettabile.

³ La durata standard è stata scelta sulla base dei valori medi della durata delle vocali italiane prodotte da una parlante nativa. La durata lunga è stata scelta sulla base della media di produzioni reali della vocale AE /ɑ:/ di una parlante nativa dell'AE, giudicate come estremamente rappresentative da un gruppo di parlanti nativi. La durata intermedia è stata scelta di conseguenza.

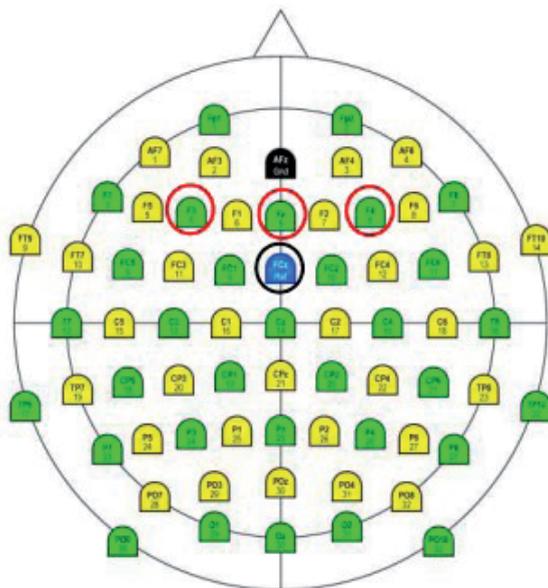
In una singola sequenza, gli stimoli standard ($p = 50\%$) sono stati presentati in maniera alternata a ciascuno dei 9 devianti ($p = 5\%$). La presentazione dei devianti è stata pseudo-randomizzata così che ogni deviante non fosse apparso mai in successione (Dev1 – St – Dev2 – St – Dev3 – St – Dev4). Inoltre, anche due devianti contenenti la stessa vocale non sono mai apparsi in successione così come due devianti aventi la stessa durata (Dev_α_180ms – St – Dev_o_255ms – St – Dev_λ_218ms).

I 9 devianti sono stati presentati in maniera equiprobabile in una serie di 18 stimoli. Gli stimoli sono stati presentati in 6 sequenze da 6 minuti circa e ogni sequenza è cominciata con la presentazione di 5 standard, non inclusi nell'analisi finale. Ogni deviante è stato presentato per un totale di 162 volte (1456 devianti totali e 1456 standard). Il SOA (*stimulu onset asynchrony*) è pari a 800ms. Il tempo totale di registrazione è stato di 38 minuti circa.

2.3.1 Registrazione EEG e analisi dei dati

Durante le sessioni di registrazione dei dati elettroencefalografici (EEG), i soggetti sedevano su una comoda sedia in una stanza schermata acusticamente. È stato detto loro di ignorare gli stimoli uditivi e di prestare attenzione ad un film senza audio scelto da loro.

Figura 2 - Configurazione degli elettrodi su una cuffia a 64 canali. In nero è evidenziato l'elettrodo utilizzato come referenza (FCz) e in rosso gli elettrodi frontali F3, Fz e F4 sui cui sono state svolte le analisi statistiche, poiché le ampiezze maggiori della MMN si registrano nella zona fronto-centrale dello scalpo



effettuata ad una frequenza di campionamento di 500Hz, è stato applicato un filtro passa-banda 0.01-70 Hz. I movimenti oculari orizzontali e verticali sono stati monitorati, rispettivamente, con F7 e FP1. Dopo la registrazione, tutti gli elettrodi sono stati riferiti al naso e l'elettrodo FCz è stato utilizzato come attivo e non come referenza. I dati grezzi sono stati segmentati in epoche di 600ms, con un pre-stimolo di 100ms relativo all'onset della pseudo-parola. È stata quindi applicata una correzione oculare (Gratton, Coles, 1983) ed i trial contenenti battiti delle palpebre sono stati esclusi dall'analisi, così come i trial con artefatti muscolari o altro rumore (rimozione degli artefatti +/- 150 μ V). La correzione della baseline è stata applicata nell'intervallo pre-stimolo di 100ms. I primi cinque stimoli standard presentati all'inizio di ogni sequenza sono stati esclusi dall'operazione di *averaging*. Questa operazione consiste nel fare la media delle risposte cerebrali a ciascun tipo di stimolo, in questo caso agli stimoli standard e agli stimoli devianti, per ottenere così un'unica onda corrispondente al potenziale evocato dallo stimolo stesso. In questo modo, il potenziale correlato all'evento (ERP, *event related potential*), cioè la presentazione dello stimolo, emerge dalla registrazione elettroencefalografica continua. I dati sono stati filtrati a 1-20Hz e le MMN sono state ottenute per ciascun soggetto sottraendo dalle onde in risposta a ciascuno dei 9 stimoli devianti le onde in risposta allo stimolo standard L1 /a/. L'ampiezza delle 9 MMN è stata analizzata nell'elettrodo centrale Fz in un intervallo di 20ms intorno al picco massimo, in due finestre temporali differenti, chiamate d'ora in poi *range*, calcolate dall'inizio della/e deviazione/i, ovvero la prima deviazione relativa alla differenza di valori formantici (100-250ms; *range* 1, R1) e la seconda deviazione relativa all'inizio della deviazione di durata (280-430 ms; *range* 2, R2).

Gli elettrodi presi in considerazione per l'analisi statistica sono gli elettrodi frontali F3 (emisfero sinistro), Fz (linea mediana) e F4 (emisfero destro), poiché la maggiore ampiezza della MMN si registra nella zona fronto-centrale dello scalpo (Näätänen, Paavilainen, Rinne & Alho, 2007).

È stata eseguita un'Analisi della Varianza a Misure Ripetute con Vocale (/a/, /Λ/, /o/) x Range (R1, R2) x Durata (standard 180ms, intermedia 218ms, lunga 255ms) x Lateralità (F3, Fz, F4). La correzione Greenhouse-Geisser è stata applicata quando necessario. Il post hoc utilizzato è quello Bonferroni.

3. Risultati

L'analisi statistica ha evidenziato una significatività per il fattore Vocale [$F(2,22) = 5,356$ $p < 0,05$]. Il post hoc, tuttavia, mostra una tendenza alla significatività ($p = 0,07$) nel confronto fra le vocali L2 /a/ vs. L1 /o/, la cui differenza in ampiezza viene comunque confermata dall'ispezione degli intervalli di confidenza (I.C.) al 95% delle ampiezze delle due vocali: l'ampiezza della vocale L2 /a/ (-0,67 μ V) è (appena) al di fuori dell'I.C. della vocale L1 /o/ (-1,34 μ V... -0,68 μ V), e l'ampiezza media della vocale L1 /o/ (-1,05 μ V) è al di fuori dell'I.C. della vocale L2 /a/ (-0,91 μ V... -0,43 μ V). Pertanto, si può ipotizzare che la vocale L2 /a/ abbia elicitato una MMN minore rispetto alla vocale L1 /o/. L'ampiezza della componente elicitata dalla vocale L2 /Λ/, invece, risulta essere equi-

valente sia a quella della vocale nativa /o/ che a quella della vocale L2 /ɑ/. Tali risultati sembrano suggerire che la vocale L2 /ɑ/ sia stata la più difficile da discriminare, la vocale L2 /ʌ/ abbia generato un livello intermedio di difficoltà e che la vocale L1 /o/ sia stata la più semplice da discriminare. Questo è in linea con la nostra ipotesi relativa alle sole differenze formantiche fra le vocali per cui la vocale L2 simile avrebbe elicitato MMN minori rispetto a quella nativa, così come quella nativa avrebbe elicitato MMN maggiori rispetto alla vocale L2 nuova.

Anche il fattore Durata è risultato essere significativo [$F(2,22) = 7,184$ $p < 0,05$] in quanto le ampiezze elicitate dalle vocali aventi durata lunga sono maggiori rispetto alle ampiezze delle vocali aventi durata pari allo stimolo standard e durata intermedia ($p < 0,05$). Questo risultato confermerebbe che il tratto di durata possa essere elaborato preattentivamente anche in parlanti che, almeno per quanto riguarda le vocali, fanno riferimento ad un sistema *quality*.

Non è stata riscontrata significatività per l'interazione Vocale x Durata ($p > 0,05$) è ciò sembra suggerire che la durata sia elaborata preattentivamente indipendentemente dalla vocale.

Figura 3 - Grand average ($N=13$) dell'onda di differenza (MMN) nell'elettrodo Fz delle vocali L2 simile /ɑ/, L2 nuova /ʌ/ e L1 /o/ delle durate standard (180ms), intermedia (218ms) e lunga (255ms) nel R1 (100-250ms) e R2 (280-430ms). Le mappe di voltaggio sono riferite al picco massimo di ampiezza delle MMN

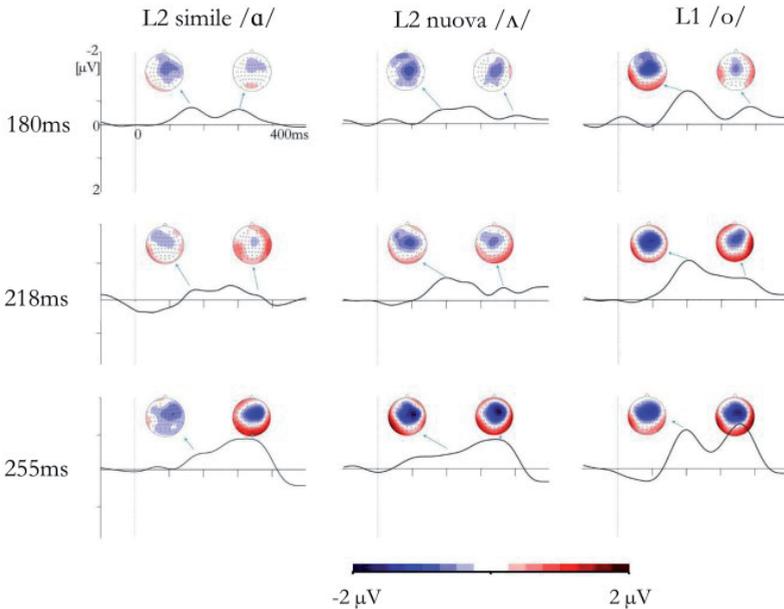


Tabella 1 - Valori medi delle ampiezze in μV (deviazione standard fra parentesi) dei picchi delle MMN in Fz nei due range: R1 (100-250) e R2 (280-430) per le vocali di durata standard (180ms), intermedia (218ms) e lunga (255ms)

	R1			R2		
	180	218	255	180	218	255
a	-0,84 (0,70)	-0,62 (0,45)	-0,81 (0,80)	-0,67 (0,58)	-0,67 (0,47)	-1,24 (0,91)
Λ	-0,86 (0,64)	-0,92 (0,56)	-1,07 (0,62)	-0,56 (0,62)	-0,94 (0,61)	-1,51 (0,90)
o	-1,21 (0,76)	-1,36 (0,85)	-1,36 (1,18)	-0,80 (0,62)	-1,17 (0,99)	-1,61 (0,99)

Il fattore Lateralità [$F(2,22) = 28,552$ $p < 0,05$] deve la sua significatività alla ampiezza minori nell'elettrodo di sinistra (F3) rispetto ai siti centrale e destro ($p < 0,05$). La distribuzione delle componenti mostra, quindi, una minore attivazione nell'emisfero sinistro.

L'interazione Range x Durata [$F(2,22) = 7,923$ $p < 0,05$] evidenzia che le vocali con durata standard (devianti singoli) hanno ampiezza maggiore nel R1 rispetto al R2 ($p < 0,05$), mentre nel R2 le vocali di durata lunga (devianti doppi) hanno ampiezza maggiore delle vocali di durata standard e intermedia (devianti doppi) ($p < 0,05$). Pertanto, coerentemente con la durata degli stimoli, si evince la presenza di una componente anche nel R2 per i devianti doppi di durata lunga (Figura 2).

L'interazione Durata x Lateralità x Range [$F(4,48) = 3,119$ $p < 0,05$] mostra che, nel R2, nell'elettrodo centrale (Fz) le vocali con durata lunga elicitano MMN maggiori rispetto alle vocali di durata standard e intermedia ($p < 0,05$) mentre, nel sito di destra (elettrodo F4), le ampiezze delle vocali di durata intermedia eguagliano quelle di durata lunga ($p > 0,05$).

Inoltre, in entrambe le finestre temporali, le componenti hanno un'ampiezza minore a sinistra rispetto al centro e a destra ($p < 0,05$) indipendentemente dalla durata e, nel R2, la durata intermedia e lunga elicitano ampiezza maggiori nel sito centrale-destro rispetto al sito sinistro ($p < 0,05$). Ciò sembra suggerire che l'aumento di durata nei devianti doppi, che sia intermedio o lungo, viene elaborato nel sito di destra.

4. Discussioni e conclusioni

Il presente lavoro ha come obiettivo lo studio dell'elaborazione preattentiva del tratto di durata vocalica in parlanti nativi dell'italiano (varietà di italiano salentino), il cui sistema fonologico prevede la presenza del tratto di durata sebbene esclusivamente per le consonanti. È stata scelta la vocale /a/ dell'italiano come standard e sono state comparate le ampiezze delle risposte MMN (componente indice dell'elaborazione e discriminazione preattentiva di stimoli linguistici) elicitate da vocali in-

glesì che vengono categorizzate con modalità differenti rispetto alle vocali italiane, ovvero la vocale L2 *simile* /ɑ/ e la vocale L2 *nuova* /ʌ/. A queste vocali inglesi, è stata aggiunta anche la vocale nativa /o/ come controllo. Le tre vocali devianti, inoltre, hanno durate differenti (standard, intermedia e lunga) per verificare se l'elaborazione della durata sia correlata con la modalità di categorizzazione dei suoni L2 rispetto a quelli L1, come dimostrato in studi precedenti (e.g., Nenonen, 2005; Lipski et al., 2012; Chládková et al., 2013). Inoltre, gli stimoli devianti che differiscono dallo stimolo standard sia per frequenze formantiche che per durata sono stati utilizzati per osservare come i devianti doppi, *double deviants*, vengano elaborati in presenza di devianti singoli in un paradigma *multi-feature* (Althen et al., 2016), ovvero se come stimolo unico o eventi distinti.

I risultati hanno dimostrato, in linea con le ipotesi iniziali relative alle sole differenze formantiche, che le vocali L2 /ɑ/ e /ʌ/ e la vocale L1 /o/ (controllo) elicitano MMN differenti in relazione alle modalità di categorizzazione rispetto al fonema nativo /a/ (Escudero, Sisinni & Grimaldi, 2014; Sisinni, Escudero & Grimaldi, 2013). Infatti, la vocale L2 simile /ɑ/ ha elicitato la componente con ampiezza minore, la vocale L2 nuova ha elicitato la componente con ampiezza intermedia mentre la vocale nativa /o/ ha elicitato la MMN con ampiezza maggiore. In effetti, la discriminazione di suoni simili, o *whitin-category*, sembra essere più difficile rispetto a suoni differenti o nuovi, o *across-category*, sia per bambini (Dehaene-Lambertz, Baillet, 1998) che per adulti (Winkler, Kujala, Tiitinen, Sivonen, Alku, Lehtokoski & Näätänen, 1999). I risultati del presente lavoro sono, quindi, in linea con gli studi precedenti in quanto la vocale L2 simile /ɑ/ può essere considerata come *whitin-category* rispetto alla vocale nativa /a/, pertanto ha elicitato MMN minori rispetto alle vocali *across-category* L2 /ʌ/ e L1 /o/. La MMN intermedia per la vocale L2 /ʌ/, a sua volta minore della MMN elicitata dalla vocale nativa /o/, è imputabile alla sua categorizzazione modale: sebbene non sia categorizzata in maniera univoca con nessuna vocale L1, la sua categorizzazione modale è con la vocale nativa /a/ (ad esempio, in Escudero et al., 2014, viene categorizzata come L1 /a/ nel 68% dei casi e come L1 /o/ nel 32%). Questa sovrapposizione percettiva fra la vocale L2 e quella L1 spiegherebbe la discriminazione intermedia (Van Leussen, Escudero, 2015; Tyler, Best, Faber & Levitt, 2014 per le previsioni sulla discriminazione di contrasti non nativi con parziale sovrapposizione percettiva).

Relativamente all'effetto della durata, le componenti maggiori sono state elicitate dalla durata lunga (255ms), più ampie rispetto a quelle elicitate dalla durata intermedia (218ms) e standard (180ms), indipendentemente dalla vocale. Ciò implicherebbe che la durata (lunga) viene elaborata a prescindere da come il suono venga categorizzato rispetto al sistema nativo. Questo risultato non è in linea con gli studi precedenti che hanno osservato una modulazione dell'elaborazione della durata relativa ai rapporti di similarità/differenza dei suoni L1/L2. Nel presente lavoro non si osserva l'azione inibitoria del sistema nativo quando percepisce la vocale L2 simile e l'azione opposta quando percepisce la vocale nuova (Nenonen et al., 2005; Chládková et al., 2013). I risultati qui ottenuti sembrano suggerire che

l'elaborazione della durata avvenga per tutte le vocali, anche per quella nativa /o/, e che l'elaborazione riguardi solo la variazione maggiore da noi proposta (durata lunga rispetto allo standard) e non quella intermedia (durata intermedia rispetto allo standard). Tuttavia, si tenga presente che in questo lavoro non è stato utilizzato un paradigma *oddball* con stimoli standard e devianti della L2 che differiscono fra loro solo in durata. Come stimolo standard è stato utilizzato uno stimolo nativo e come devianti, doppi devianti contenenti le vocali L2 e la vocale nativa. I paradigmi sperimentali, quindi, non sono strettamente comparabili. Conseguentemente, anche i risultati possono non essere facilmente paragonabili fra loro.

La difficoltà di comparazione emerge anche con il lavoro di Ylinen et al. (2010). Come nel presente lavoro, anche Ylinen e colleghi hanno utilizzato doppi devianti ma non in paradigmi *multi-feature* e non in presenza di devianti singoli. Ylinen et al. (2010) hanno trovato una MMN unica in risposta ai doppi devianti, diversamente dalle nostre MMN che presentano due picchi (come si discuterà anche in seguito). Studi successivi, pertanto, si rendono necessari per omologare quanto più possibile le procedure sperimentali per avere una visione più coerente dei risultati.

Il fatto che sia la durata lunga ad elicitar MMN maggiori e non la durata intermedia sembra poter suggerire l'esistenza di una soglia di salienza della durata. I nostri dati non ci permettono di identificare una soglia ben precisa, sebbene sembrino suggerire che una soglia minore o uguale ai 38 ms (differenza fra durata standard e intermedia) non abbia un effetto significativo, al contrario di una soglia maggiore o uguale a 75ms (differenza fra durata standard e lunga). Nell'emisfero destro, tuttavia, le ampiezze maggiori sono elicitate anche dalla durata intermedia, e non solo da quella lunga: ciò suggerirebbe che l'emisfero destro è più sensibile alle differenze di durata, anche con durate "intermedie". L'elaborazione a destra della durata intermedia e lunga è in linea con i risultati ottenuti da Nenonen et al. (2005) che ipotizzano che la MMN in risposta a stimoli caratterizzati dal tratto della durata, elicitata in parlanti di sistemi qualitativi, non può riflettere un'elaborazione fonologica ma, piuttosto, un'elaborazione acustica, come accade per suoni non linguistici (Giard, Lavikahen, Reinikainen, Perrin, Bertrand, Pernier & Näätänen, 1995; Paavilainen, Alho, Reinikainen, Sams & Naatanen, 1991). Nenonen et al. (2005) suggeriscono un'ulteriore ipotesi, ovvero che la maggiore attivazione nell'emisfero destro possa essere ricondotta ad un'attivazione frontale che maschera l'attivazione temporale sinistra (Giard, Perrin, Pernier & Bouchet, 1990; Opitz, Rinne, Mecklinger, von Cramon & Schröger, 2002; Rinne, Alho, Ilmoniemi, Virtanen & Näätänen, 2000). Tuttavia, considerando che i parlanti da noi analizzati non posseggono il tratto fonologico di durata vocalica, si può ipotizzare che si tratti di un'elaborazione di tipo meramente acustico (Näätänen, 2001). Alternativamente, non si può escludere che la durata vocalica sia elaborata da un punto di vista prosodico.

Il ricorso ai doppi devianti ci ha permesso di osservare come due deviazioni all'interno della TWI siano elaborate. Studi precedenti (e.g., Czigler, Winkler, 1996) hanno rilevato che due deviazioni presenti in uno stesso stimolo deviante sono elaborate come un unico evento anche quando presenti devianti singoli, la cui

presenza dovrebbe causare l'elaborazione della seconda deviazione che, in tale contesto, sarebbe rilevante. Dai nostri risultati parrebbe che le due deviazioni vengano elaborate come eventi sequenziali e non come un evento unico frutto di integrazione di processi paralleli: sia nel R1 che nel R2, per tutte e tre le durate, è stata osservata la presenza di una componente. Tuttavia, nel R2, per le vocali con durata standard, vi è una componente con ampiezze minori rispetto al R1, mentre per le vocali con durata intermedia e lunga, le componenti hanno pari ampiezza in entrambi i range. Osservando la morfologia delle componenti (Figura 2), inoltre, si evince chiaramente la presenza di due picchi. Ciò è particolarmente evidente per la vocale L1 /o/ di durata 255ms, ma anche per le vocali L2 aventi durata lunga, che nel R2 presentano un picco relativo all'elaborazione della seconda deviazione, la durata appunto (mentre nel R1 presentano un picco più basso).

I nostri dati, pertanto, suggeriscono che due deviazioni, sebbene veicolate dallo stesso stimolo, possono essere elaborate come eventi distinti e sequenziali. La presenza di devianti singoli (con durata pari a quello dello standard) fungerebbe da contesto informativo (Sussman, Winkler, 2001): la seconda deviazione presente nei devianti doppi veicolerebbe una nuova informazione che verrebbe, conseguentemente, elaborata.

Il presente studio differisce, quindi da, ad es., Czigler, Winkler (1996) ed è in linea con gli studi di Oceak et al. (2008) e Althen et al. (2016) che hanno osservato un'elaborazione sequenziale delle due deviazioni presenti nei devianti doppi. In entrambi questi studi i doppi devianti differiscono dallo standard per "frequenza+intensità"; nel presente studio, invece, differiscono per "frequenza+durata", deviazioni per le quali altri studi (Wolff, Schröger, 2001; Paavilainen et al., 2003) hanno osservato un solo picco con stimoli non linguistici.

Non si può escludere che la natura degli stimoli qui utilizzati abbia influito sul presente risultato. È necessario, infatti, interpretare i risultati ottenuti in una prospettiva prettamente linguistica: i nostri soggetti, parlanti nativi dell'italiano, *quantity system* per ciò che concerne le sole consonanti, e non per le vocali, sembrano aver elaborato i valori formantici prima, determinanti per il processo di categorizzazione e, solo successivamente, la durata. Ciò suggerirebbe che l'elaborazione della durata vocalica nei parlanti italiani non avviene come nei parlanti nativi di *quantity systems* "puri", per i quali si ipotizza l'esistenza di due meccanismi neurali indipendenti e paralleli, uno per la dimensione formantica e l'altro per la durata (Ylinen et al., 2005). Ulteriori ricerche sono necessarie per verificare se l'elaborazione della durata consonantica sia in linea con quella dei parlanti *quantity*, avendo valenza fonologica per i parlanti italiani.

In conclusione, l'uso di suoni L2 simili e nuovi, così come l'uso di devianti doppi, sembra suggerire che quando i parlanti nativi dell'italiano (*quantity system* per le consonanti, ma non per le vocali) percepiscono suoni vocalici, elaborano i valori formantici e la durata come eventi separati e sequenziali: in primo luogo elaborano i valori formantici, fondamentali per il processo di categorizzazione e, solo successivamente, elaborano la loro durata, indipendentemente dalla vocale che percepiscono.

Riferimenti bibliografici

- ALTHEN, H., HUOTILAINEN, M., GRIMM, S. & ESCERA, C. (2016). Middle latency response correlates of single and double deviant stimuli in a multi-feature paradigm. In *Clinical Neurophysiology*, 127(1), 388-396.
- BEST, C.T., TYLER, M.D. (2007). Nonnative and second language speech perception: Commonalities and complementarities. In MUNRO, M.J., BOHN, O.-S. (Eds), *Second Language speech learning: the role of language experience in speech perception and production*. Amsterdam: John Benjamins, 13-34.
- BOGACKA, A. (2004). On the perception of English high vowels by Polish learners of English. In DASKALAKI, E., KATSOS, N., MAVROGIORGOS, M. & REEVE, M. (Eds.), *CamLing 2004: Proceedings of the University of Cambridge second postgraduate conference in language research*. Cambridge: Cambridge University Press, 43-50.
- BOHN, O.-S. (1995). Cross language speech production in adults: First language transfer doesn't tell it all. In STRANGE, W. (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press, 279-304.
- CEBRIAN, J. (2006). Experience and the use of duration in the categorization of L2 vowels. In *Journal of Phonetics*, 34, 372-387.
- CHLÁDKOVÁ, K., ESCUDERO, P. & LIPSKI, S.C. (2013). Pre-attentive sensitivity to vowel duration reveals native phonology and predicts learning of second-language sounds. In *Brain and language*, 126(3) 243-252.
- CZIGLER, I., WINKLER, I. (1996). Preattentive auditory change detection relies on unitary sensory memory representations. In *NeuroReport*, 7(15-17), 2413-2418.
- DEHAENE-LAMBERTZ, G., BAILLET, S. (1998). A phonological representation in the infant brain. In *NeuroReport*, 9,8, 1885-1888.
- ESCUDEO, P., BENDERS, T., LIPSKI, S.C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. In *Journal of Phonetics*, 37(4), 452-465.
- ESCUDEO, P., SISINNI, B. & GRIMALDI, M. (2014). The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. In *The Journal of the Acoustical Society of America*, 135(3), 1577-1584.
- FLEGE, J.E., BOHN, O.-S. & JANG, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. In *Journal of Phonetics*, 25, 437-470.
- GIARD, M.H., LAVIKAHEN, J., REINIKAINEN, K., PERRIN, F., BERTRAND, O., PERNIER, J. & NÄÄTÄNEN, R. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: an event-related potential and dipole-model analysis. In *Journal of Cognitive Neuroscience*, 7(2), 133-143.
- GIARD, M.H., PERRIN, F., PERNIER, J. & BOUCHET, P. (1990). Brain Generators Implicated in the Processing of Auditory Stimulus Deviance: A Topographic Event-Related Potential Study. In *Psychophysiology*, 27(6), 627-640.
- HORVÁTH, J., CZIGLER, I., WINKLER, I. & TEDER-SÄLEJÄRVI, W.A. (2007). The temporal window of integration in elderly and young adults. In *Neurobiology of aging*, 28(6), 964-975.

- IVERSON, P., KUHL, P.K., AKAHANE-YAMADA, R., DIESCH, E., TOHKURA, Y.I., KETTERMANN, A. & SIEBERT, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. In *Cognition*, 87(1), B47-B57.
- JARAMILLO, M., ILVONEN, T., KUJALA, T., ALKU, P., TERVANIEMI, M. & ALHO, K. (2001). Are different kinds of acoustic features processed differently for speech and non-speech sounds?. In *Cognitive Brain Research*, 459-466.
- KONDAUROVA, M., FRANCIS, A. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. In *Journal of the Acoustical Society of America*, 124, 3959-3971.
- LIPSKI, S.C., ESCUDERO, P. & BENDERS, T. (2012). Language experience modulates weighting of acoustic cues for vowel perception: An event-related potential study. In *Psychophysiology*, 49(5), 638-650.
- MORRISON, G.S. (2002). Perception of English /i/ and / I / by Japanese and Spanish listeners: Longitudinal results. In MORRISON, G.S., ZSOLDOS, L. (Eds.), *Proceedings of the north west linguistics conference 2002*. Burnaby, BC, Canada: Simon Fraser University Linguistics Graduate Student Association, 29-48.
- NÄÄTÄNEN, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). In *Psychophysiology*, 38(1), 1-21.
- NÄÄTÄNEN, R., PAAVILAINEN, P., RINNE, T. & ALHO, K. (2004). The mismatch negativity (MMN) in basic research of central auditory processing: A review. In *Clinical Neurophysiology*, 118, 2544-2590.
- NÄÄTÄNEN, R., PAKARINEN, S., RINNE, T. & TAKEGATA, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. In *Clinical Neurophysiology*, 115(1), 140-144.
- NENONEN, S., SHESTAKOVA, A., HUOTILAINEN, M. & NÄÄTÄNEN, R. (2005). Speech-sound duration processing in a second language is specific to phonetic categories. In *Brain and language*, 92(1), 26-32.
- OCEÁK, A., WINKLER, I. & SUSSMAN, E. (2008). Units of sound representation and temporal integration: A mismatch negativity study. In *Neuroscience letters*, 436(1), 85-89.
- OPITZ, B., RINNE, T., MECKLINGER, A., VON CRAMON, D.Y. & SCHRÖGER, E. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. In *NeuroImage*, 15, 167-174.
- PAAVILAINEN, P., ALHO, K., REINIKAINEN, K., SAMS, M. & NÄÄTÄNEN, R. (1991). Right hemisphere dominance of different mismatch negativities. In *Electroencephalography and clinical neurophysiology*, 78(6), 466-479.
- PAAVILAINEN, P., MIKKONEN, M., KILPELAINEN, M., LEHTINEN, R., SAARELA, M. & TAPOLA, L. (2003). Evidence for the different additivity of the temporal and frontal generators of mismatch negativity: a human auditory event-related potential study. In *Neuroscience Letters*, 379, 79-82.
- RAUBER, A.S., ESCUDERO, P., BION, R. & BAPTISTA, B.O. (2005). The interrelation between the perception and production of English vowels by native speakers of Brazilian Portuguese. In *Proceedings of Interspeech 2005*, 2913-2916.

- RINNE, T., ALHO, K., ILMONIEMI, R.J., VIRTANEN, J. & NÄÄTÄNEN, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. In *Neuroimage*, 12(1), 14-19.
- SISINNI, B., ESCUDERO, P. & GRIMALDI, M. (2013). Salento Italian listeners' perception of American English vowels. In *Proceedings of Interspeech 2013*, 2091-94.
- SUSSMAN, E., WINKLER, I., RITTER, W., ALHO, K. & NÄÄTÄNEN, R. (1999). Temporal integration of auditory stimulus deviance as reflected by the mismatch negativity. In *Neuroscience letters*, 264(1), 161-164.
- TYLER, M.D., BEST, C.T., FABER, A. & LEVITT, A.G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. In *Phonetica*, 71(1), 4-21.
- VAN LEUSSEN, J.-W., ESCUDERO, P. (2015). Learning to perceive and recognize a second language: the L2LP model revised. In *Frontiers in Psychology*, 6, 1-12.
- WANG, W., DATTA, H. & SUSSMAN, E. (2005). The development of the length of the temporal window of integration for rapidly presented auditory information as indexed by MMN. In *Clinical neurophysiology*, 116(7), 1695-1706.
- WINKLER, I., KUJALA, T., TIITINEN, H., SIVONEN, P., ALKU, P., LEHTOKOSKI, A. & NÄÄTÄNEN, R. (1999). Brain responses reveal the learning of foreign language phonemes. In *Psychophysiology*, 36(5), 638-642.
- WOLFF, C., SCHRÖGER, E. (2001). Human pre-attentive auditory change-detection with single, double, and triple deviations as revealed by mismatch negativity additivity. In *Neuroscience Letters*, 311, 37-40.
- YLINEN, S., HOUTILAINEN, M. & NÄÄTÄNEN, R. (2005). Phoneme quality and quantity are processed independently in the human brain. In *Neuroreport*, 1857-1860.
- YLINEN, S., UTHER, M., LATVALA, A., VEPSÄLÄINEN, S., IVERSON, P., AKAHANE-YAMADA, R. & NÄÄTÄNEN, R. (2010). Training the brain to weight speech cues differently: A study of Finnish second-language users of English. In *Journal of Cognitive Neuroscience*, 22(6), 1319-1332.