Short Title:

Learning two languages from birth shapes pre-attentive processing of speech

Full Title:

Learning two languages from birth shapes pre-attentive processing of vowel categories: Electrophysiological correlates of vowel discrimination in monolinguals and simultaneous bilinguals

Monika Molnar [1,], Linda Polka [2,3], Shari Baum[2,3], and Karsten Steinhauer[2,3]

[1]Basque Center on Cognition, Brain, and Language (BCBL), Paseo Mikeletegi 69, Donostia 20009, Spain. Tel**:** +34 943 309 300 Fax: +34 943 309 052

[2]McGill University, School of Communication Sciences and Disorders, 1266 Pine Ave W, Montreal, H3G 1A8 Canada. Tel.: 514-398-4137 Fax: 514-398-8123

[3]Centre for Research on Language, Mind, and Brain (CRLMB), 640 de la Montagne, Montreal QC H3G 2A8 Canada. Tel.: 514-398-4137 Fax: 514-398-8123

Corresponding author: m.molnar@bcbl.eu

Abstract

Using event-related brain potentials (ERPs), we measured pre-attentive processing involved in native vowel perception as reflected by the mismatch negativity (MMN) in monolingual and simultaneous bilingual (SB) users of Canadian English and Canadian French in response to various pairings of four vowels: English /u/, French /u/, French /y/, and a control /y/. The monolingual listeners exhibited a discrimination pattern that was shaped by their native language experience. The SB listeners, on the other hand, exhibited a MMN pattern that was distinct from both monolingual listener groups, suggesting that the SB pre-attentive system is tuned to access sub-phonemic detail with respect to both input languages, including detail that is not readily accessed by either of their monolingual peers. Additionally, SBs exhibited sensitivity to language context generated by the standard vowel in the MMN paradigm. The automatic access to fine phonetic detail may aid SB listeners to rapidly adjust their perception to the variable listening conditions that they frequently encounter.

Keywords: *MMN, vowel perception, bilingualism, language context*

1.1. Introduction

The phonological organization of bilinguals has been of great interest to speech perception and production research in recent decades. Most studies have focused on the abilities of sequential bilinguals, who acquired their first language (L1) from birth, and started to learn a second language (L2) at later stages of their lives (Best, 1995; Flege, 1995). One of the general conclusions has been that sequential bilinguals are able to achieve native monolingual-like phonological performance in the L2, if the bilinguals' onset of L2 acquisition is close to the onset of L1 acquisition (birth) and/or their L2 proficiency is high (e.g., Bongaerts, Mennen, and van der Silk, 2000; Guion, 2003; Piske et al., 2002; Flege et al., 1999, 2003; Mackay et al, 2001; Perani et al., 1998; Yamada, 1995; but see: Pallier and Sebastián-Gallés , 1997). By contrasting sequential bilinguals' L2 abilities with the performance of native monolinguals, it is inferred, however, that fully successful second language acquisition and an intact bilingual phonological organization require native monolingual-like performance. This phenomenon has been defined as the monolingual or *fractional view* on bilingualism (Grosjean, 1998).

Based on this fractional view, one might expect that simultaneous bilingual (SB) individuals — those who developed two languages within the same timeframe from birth and are therefore native speakers of two languages — would exhibit monolingual-like abilities within both of their native languages. However, findings of behavioral studies (e.g., Caramazza et al., 1973; Guion, 2003; Mack, 1989; Sebastián-Gallés , Echevirra, and Bosch, 2005; Sundara, Polka & Baum, 2006; Sundara and Polka, 2008) seem to converge toward the conclusion that speech perception and production patterns of SB individuals do not completely match the patterns observed in their monolingual peers.

For instance, in terms of consonants, SB adults accommodate additional phonetic detail that is not consistently produced and perceived by either of their monolingual peers (e.g., Sundara, 2005; Sundara et al., 2006; Sundara and Polka, 2008). As well, there is evidence that SB adults' production patterns are more streamlined than their monolingual peers; SB adults appear to selectively omit free variant forms that fail to enhance phonetic category differences in either of their languages (MacLeod and Stoel-Gammon, 2005; Sundara et al., 2006). Similar findings have emerged for vowels. For instance, even though simultaneous bilingual speakers of Quechua and Spanish produce vowel categories of the two languages similar to those of the monolingual speakers, the bilingual productions (in terms of phonetic detail) differ from the monolingual ones. The bilingual productions of the Quechua vowels are raised upward along the F1 axis in the acoustic/phonetic vowel space, as compared to monolingual productions of the same vowels (Guion, 2003). This systemic manipulation of the bilingual vowel categories probably serves to sustain separate vowels across and within the two languages.

Considering SB adults' processing of cross-language vowels and consonants, SB adults do not appear to have a dual set of entirely monolingual-like speech processing abilities, even though they are native speakers of both languages under investigation. Rather, it seems that SBs represent a group of language users distinct from monolingual users of both of their languages — a conclusion that is in line with the so-called *holistic view* of bilingualism (Grosjean, 1998).

Surprisingly, the neural correlates of SB language processing that contribute to such unique behavioral patterns have only rarely been investigated (e.g., Kim et al, 1997; Klein et al., 1995; Peltola et al., 2012; Perani et al., 1998). With respect to phonological

processing, the most common tool to measure vowel or consonant processing in monolinguals and sequential bilinguals has been a specific type of neurophysiological measure, the mismatch negativity or MMN (e.g., Näätänen et al., 1997; Sierra-Garcia, 2012; Winkler et al., 1999). However, this measure has been only recently applied in a SB population (Peltola et al., 2012). The MMN is an early event-related brain potential (ERP) component sensitive to change detection and discrimination in auditory stimulation (Näätänen and Picton, 1987). The classic method for eliciting MMN is via the oddball paradigm (for review: Näätänen et al, 2007), which involves presenting a sequence of identical stimuli (referred to as standards) in which rare irregular stimuli (referred to as deviants) appear approximately 5 to 30 percent of the time. Participants normally listen to such stimuli streams passively without paying direct attention, while watching silent videos or reading books. In the corresponding ERPs, both standards and deviants elicit an early negativity (around 100 ms after stimulus onset), called N100 or N1, followed by a positive peak (around 200 ms after stimulus onset) called P200 or P2. Importantly, only the deviant stimulus evokes an additional MMN component (e.g., a negative wave with peak latencies between 150 and 300 ms). This MMN effect is most evident in the ERP *difference wave*, when the ERP response to the standard is subtracted from the response to the deviant stimuli.

It has been demonstrated that the phonemic status of the speech sounds might enhance the amplitude or decrease the latency of the MMN as compared to components elicited by non-phonemic contrasts, indicating that language experience shapes the early pre-attentive processing of speech sounds (e.g., Näätänen et al., 1997; Dehaene-Lambertz, 1997; Sharma and Dorman, 1999; Winkler et al, 1999). For example, Winkler

and colleagues (1999) investigated the vowel discrimination abilities of native Finnish and native Hungarian participants in response to Finnish and Hungarian vowel contrasts that either represent an across-phonemic category change in both languages, or contrasts that represent a phonemic change in Finnish, but a within-category vowel contrast in Hungarian (and vice versa). Both the within-category and the across-phonemic category changes elicited an MMN in both language groups; however, the participants exhibited an earlier latency for the across- than the within-category vowel contrasts.

To explain the enhancement of the MMN (in terms of larger amplitude or shorter latency) in the case of phonemic (versus non-phonemic) contrasts, Näätänen et al. (1997) suggested that traces of categorical prototypes in *cortical long-term memory* (auditory cortex) serve as "recognition patterns" for native phonemes, facilitating the correct perception of speech sounds. Thus, if a given vowel happens to be a prototype of a phonemic category, its percept is reinforced by the corresponding recognition pattern stored in long-term memory. During the presentation of deviant sounds that are prototypical (or phonemic), this effect elicits a larger MMN when contrasted with non-phonemic deviants. Furthermore, it has been generally concluded that phoneme representations are based on language-specific cortical long-term memory traces, which begin to develop in the first year of life (e.g., Peltola et al., 2003).

A more recent study by Peltola and colleagues (2012) investigated the neural correlates of speech perception in Finnish-Swedish simultaneous (who began to learn both languages during infancy) and sequential bilinguals (who began to learn Finnish during infancy and Swedish later in life), and reported differences in the neural organization of vowel discrimination across the two bilingual groups. MMN responses

elicited by two closed rounded vowels (/y/, and /ʉ/) were investigated. Importantly, the two vowels cross a vowel category boundary in Finnish (phonemic difference), but the same vowel pair represents a within-category change in Swedish (allophonic difference). Each bilingual participant was tested twice on different days using the same oddball paradigm: once in a Finnish linguistic context, and once in a Swedish linguistic context.

Peltola and colleagues (2012) reported two types of effects in the MMN responses. First, only the MMN amplitudes elicited by sequential but not by simultaneous bilinguals were affected by the linguistic context. Sequential bilinguals exhibited a larger MMN response in the Finnish (L2) context (in which the vowels represented two different phonemes), and showed no significant MMN response in the Swedish (L1) context (in which the vowels represented an allophonic change); yet, the MMN amplitudes exhibited by simultaneous bilinguals were not affected by linguistic context. Second, the latency of the MMN was slightly influenced by the linguistic context in the simultaneous bilingual group, as the ERP response was somewhat faster in the Finnish context (phonemic change) as compared to the Swedish context (allophonic change). According to the authors' interpretation, the results related to amplitude changes clearly illustrate that simultaneous bilinguals, but not sequential bilinguals, rely on a shared phonological space across their languages ("one-store model"), as the MMNs of simultaneous bilinguals were not affected by linguistic context.

Similar results have been reported for Spanish-English sequential bilinguals' discrimination of consonants (Garcia-Sierra et al., 2012; but see Winkler et al., 2003). When sequential bilingual participants passively attended to phonemic and non-phonemic

VOT changes in English and Spanish language contexts, their MMN patterns suggested that pre-attentive auditory change detection is shaped by linguistic context.

1.2 The present study

The present study was designed to explore the MMN responses to both acoustically similar and distinct cross-language vowel category representations in Canadian English and Canadian French SB adults and in monolingual speakers of these languages. As opposed to previous investigations with bilinguals (Garcia-Sierra et al., 2012; Peltola et al., 2012), we were particularly interested in the processing of cross-language speech sounds when no specific external linguistic context is established by using a language-specific task during the passive oddball paradigm (e.g., to establish language context, participants could read magazines in the context-appropriate language).

The advantage of obtaining MMNs without establishing an external language context is that identical testing conditions across monolinguals and bilinguals are provided. The presence of language context effects in *monolingual* settings tends to minimize actual differences between monolinguals and bilinguals (e.g., Garcia-Sierra et al., 2012) making it difficult to establish the degree to which learning two languages from birth shapes the perceptual processing of speech. On the other hand, a potentially more revealing '*bilingual mode*' (when both languages are activated) can only be applied, naturally, with bilinguals but not with monolinguals. Thus, using a measure such as the MMN (without language context manipulations) that is sensitive to subtle differences in phoneme perception but not necessarily susceptible to external context effects may fill an important methodological gap for understanding bilingual processing, because it taps into

the earliest processing level that precedes conscious perception of a sound that is unavailable for investigation by traditional behavioral methods (Sams et al, 1985; Näätänen, 1987).

Based on a previous behavioral investigation of SB and monolingual populations in our laboratory (Molnar, Polka, Baum et al., 2010), we selected three vowels to be used in the oddball paradigm (as illustrated in **Figure 1**). The first vowel (V1) was an /u/ sound corresponding to the best exemplar (or prototype) of the native Canadian French vowel (as in the French word 'douze'); V2 was an /u/ sound representing the prototypical Canadian English vowel (as in the English word 'you'); and V4 corresponded to the best /y/ vowel (as in the French word 'sucre') as identified by native Canadian French speakers. Importantly, this vowel (/y/) is non-phonemic in English. In a second step, we also selected a fourth control vowel that falls between the English /u/ and the French /y/, along an F1/F2 continuum. With the purpose of achieving a fully symmetrical design in terms of psychoacoustic differences, we chose this control vowel (V3) such that its perceptual distance (on a Bark scale) from the French /y/ (V4) was identical to the distance between the English /u/ (V2) and the French /u/ (V1).

Insert Figure 1. approximately here

In four multiple-deviant oddball paradigms (see details in the Method Section), the discrimination of all four vowels was measured resulting in 6 contrasts (V1 vs. V2, V3 vs. V4, V1 vs. V4, V2 vs. V3, V1 vs. V3, and V2 vs. V4). Two of these contrasts represent a *within-phonemic category change*: within the French [y] (V4 vs. V3), and within the English or French [u] (V1 vs. V2). The V1-V2 contrast represents allophonic

variations, and it includes vowels that are prototypical in both languages, thus similar MMN responses across all groups are expected. The V3-V4 contrast (French [y] vs. control vowel), on the other hand, represents a within-category change only for the French and SB groups. Thus different (e.g., reduced) MMN responses are expected to be elicited from the English listeners as compared to the two other groups in response to the V3-V4 vowel pair.

The rest of the four contrasts involve *across-phonemic category* changes: between (good or poor) exemplars of the French [y] and (good or poor) exemplars of the French/English [u]. Based on the language background of the participants, if a good/prototypical exemplar of a native vowel is present in the contrast, a MMN response should be present as well. In the case of the V2-V3 vowel pair (control vowel vs. English [u]), only poor exemplars of French vowels are present, therefore an MMN might be present in the English, but not in the French group. However, the relatively small acoustic difference between V2 and V3 might not be sufficient to elicit a MMN in either of the groups. Based on the fractional view of bilingualism, SBs are expected to show a pattern similar to the monolingual English group. In contrast, in keeping with the holistic view of bilingualism, SBs would be expected to show a different pattern from the monolingual groups.

The V1-V2 and V2-V4 contrasts represent two relevant pairs for the bilinguals, because both vowels within each pair belong to different languages. However, the V1-V2 also reflects within-category variations for both the English and French vowels, whereas the V2-V4 pair represents clear categorical changes across the languages. If SBs are sensitive to differences that are relevant for distinguishing phonemes across languages,

SBs are expected to show different MMN patterns (e.g., reflected in amplitude or latency) from the monolingual listeners, at least in response to the V2-V4 contrasts.

2. Method

2.1 Participants

Forty-nine right-handed participants with no history of speech, language, or hearing impairment signed an informed written consent to take part in the experiment. EEG recordings of four participants were excluded from the analysis due to technical problems (2) and poor data quality caused by artifacts (2). Participants' language background was assessed using three measures: (1) a detailed language questionnaire filled out via interview with a proficient bilingual research assistant; (2) the *Language Experience and Proficiency Questionnaire* (LEAP-Q) that was developed specifically to evaluate bilingual and multilingual individuals' language background (Marian et al., 2007); (3) a 5-minute speech sample evaluated by monolingual speakers of Canadian English (n=3) and Canadian French (n=3) using a scale from 1 to 5, where 5 represents native-like ability and 1 indicates no ability in the given language. Based on this assessment (as detailed below), 15 monolingual English speakers from English-speaking Canada (average age=25 years, 7 females), 15 monolingual French speakers from French-speaking Canada (average age=26 years, 7 females), and 15 simultaneous bilingual speakers of Canadian English and Canadian French (average age=23 years, 9 females) were included in the final analysis.

Only participants who fit the following criteria were included in the monolingual English and monolingual French groups: (1) both parents of the participant are monolingual speakers of the appropriate language; (2) the participant was educated in

monolingual school settings in the appropriate language; (3) they began formal learning of a second language in school settings not earlier than age 10; (4) they do not use a second language on a regular basis, having rated their ability in a second language with a maximum of 4 out of 10; (5) their native speech sample was rated 5 (native-like) on average.

Inclusion criteria for the SB group were the following: (1) participants had been exposed to both languages since birth, both languages are spoken in their homes, and the participants still use both languages on an everyday basis; (2) their schooling was completed, at different points, in both English and French; (3) participants rated their language abilities for both languages with a minimum of 9 out of 10 and the bilingual interviewer confirmed their self-ratings; (4) they have been living in bilingual areas of Canada, such as Montreal or the Ottawa region; (5) the monolingual raters evaluated their speech samples in both languages with a minimum average of 4.5 out of 5.

2.2 Stimuli and Procedure

Based on our previous behavioral experiment (Molnar,et al., 2010), we selected four synthesized vowels to represent a good exemplar of Canadian French /u/ (F1= 275 Hz, F2= 745 Hz), a good exemplar of Canadian English /u/ (F1= 300 Hz, F2= 979 Hz), a good exemplar of Canadian French /y/ (F1= 275Hz, F2= 2011 Hz), and a control vowel representing a non-prototypical Canadian French /y/ (F1= 300 Hz, F2=1597Hz), as illustrated in Figure 1. Each token had fixed F3 (2522 Hz) values and was 400 ms in duration including a 35 ms rise and fall time.

The vowels were presented at 65 dB HL via insert earphones (Etymotic Research) in four different experimental blocks, as illustrated in Figure 2. Each block contained a total of 1000 stimuli: 790 standard sounds and 210 deviants (70 of each deviant) presented with a 1600 ms ISI. Within each block, standard and deviant sounds were in a pseudo-random order ensuring that at least two standard sounds preceded one deviant token. All four blocks were played within one testing session in a pseudo-randomized order, counterbalanced across participants (Latin square) and across groups. During the EEG recording, participants were sitting in a comfortable armchair in an electrically shielded sound-attenuated booth watching a silent movie. The experimental sessions lasted approximately 3.5 hours including preparation time (approximately 40 minutes), and breaks (approximately 30 minutes).

In the oddball paradigm, the vowels were presented using a relatively long inter-stimulus-interval (ISI) of 1600 ms. The ISI is the duration of the silent gap between two sounds as measured from the offset of one stimulus sound to the onset of the next stimulus. EEG oddball paradigms normally employ relatively short ISIs (approximately 500 ms) in order to sufficiently strengthen the short-term memory trace for the repeated standard sounds that develops online during the experiment. The MMN is elicited when the deviant stimulus is compared against this trace and results in a mismatch. That is, the shorter the ISI, (i) the less likely it is that this trace in short-term memory will decay between two stimuli and (ii) the more standards can contribute to the trace within the relevant time window (a few seconds). In contrast, with a longer ISI, the short-term memory trace decay is more complete, and thus the sound discrimination process underlying the MMN may rely to a larger extent on the categorization of the first sound

based on the recognition pattern in long-term memory representations. In fact, behavioral speech discrimination studies (Carney et al. 1977; Pisoni, 1973; Werker and Logan, 1985) have illustrated that language-specific ('phonemic') effects, which rely on neural traces developed for long-term memory representations of speech sounds, are more obvious when the ISI is relatively long (at least 1000 ms), while a short ISI tends to primarily activate acoustic short-term memory, thus tapping into acoustic processing that may be influenced by both short-term acoustic memory and long-term memory of phonemic traces. The primary interest of the current study is to examine memory traces developed in long-term memory (since birth), which plays a major role in everyday speech processing, and to minimize the role of online developing memory traces. To achieve this, a relatively long ISI (1600 ms) was implemented in the experiment presented here.[1] Note that a pilot ERP study was conducted prior to the main experiment in order to determine whether the long 1600 ms ISI would prevent MMN effects. In this pilot study, we tested ISIs varying between 700 and 1600 ms, and found no differences in MMN amplitude for the various ISI conditions. Based on these results, we decided that the 1600 ms ISI was indeed compatible with an MMN oddball paradigm.

2.3 EEG Recording

EEG was continuously recorded (500 Hz/32 bit sampling rate; Neuroscan Synamps2 amplifier) from the scalp with cap-mounted Ag-Ag/Cl electrodes (Electro-cap International, Inc, Eaton, OH) from 20 sites on the scalp, based on the international 10-20 system of electrode placement: Fp1, Fp2, F7, F8, F3, F4, T3, T4, C3, C4, T5, T6, P3, P4,

---

[1] Note that a pilot ERP study was conducted prior to the main experiment in order to determine whether the long 1600 ms ISI would prevent MMN effects. In this pilot study, we tested ISIs varying between 700 and 1600 ms, and found no differences in MMN amplitude for the various ISI conditions. Based on these results, we decided that the 1600 ms ISI was indeed compatible with an MMN oddball paradigm.

O1, O2, Fpz, Fz, Cz, Pz, and Oz. Vertical (EOGV) and horizontal (EOGH) eye-movements were monitored by bipolar electrodes placed above and below the left eye, and at the outer corner of each eye, respectively. All electrodes were referenced against the right mastoid, and an electrode located between Fz and Fpz provided the ground. Electrode impedances were kept under 3 kOhm.

2.4 Data analysis

EEG data were analyzed using the Brain Vision Analyzer software (Brain Products GmbH, Germany), including offline band-pass filtering (.5 to 30 Hz[2]) and artifact rejection with a +/- 50 μV deviation criterion at all channels except for Fp1 and Fp2, which were clearly more affected by eye movements than the rest of the channels. Consequently, Fp1 and Fp2 were excluded from any further ERP analysis and data processing. Artifact rejection resulted in data loss within a range of 3.45% and 11.09% across participants.

ERP averages were time-locked to vowel onset and were computed separately for the standard and deviant conditions of each vowel. Only averages for standards immediately preceding a deviant stimulus were included in the analyses; therefore, approximately the same number of standards and deviants contributed to the final grand averages. Epochs (900 ms) included a 100 ms baseline starting 50 ms before and ending 50 ms after stimulus onset, as each stimulus was fully audible after 50 ms of stimulus onset only.

As illustrated in **Figure 2**, all the vowels were presented both as standards and as deviants across the four different presentation blocks, unlike in most MMN studies which

---

[2] Analyses with other band-pass settings (0.4-40 Hz, 0.4-100 Hz) were also computed. They resulted in the same findings reported here, but data included more noise.

assign the role of standard to one specific sound alone and the role of deviant to the other (remaining) sound(s) of interest. However, different sounds, due to differences in their pure physical properties, might elicit different N1 and P2 components (overlapping with the MMN) even when they are not presented in an oddball task. Therefore, if the role of standard and deviant is confounded with the different sounds used in an oddball paradigm, the MMN components might be partially influenced by the physical differences between the sounds, in addition to the oddball effect under investigation (for related discussion of ERP artifacts see also Steinhauer & Drury, 2012).

Another reason why it is crucial to employ all the sounds as both standards and deviants is to avoid any confounds with the order of stimulus presentation. It has been demonstrated in behavioral studies (Natural Referent Vowel Hypothesis; Polka and Bohn, 2003; Polka and Bohn, 2011), and even in ERP investigations (Polka et al., 2009), that the discriminability of sounds is affected by the order in which they are presented relative to each other. In addition, whether the sound presented first is more prototypical in the given language compared to the following sounds (or the other way around) has also been shown to influence discrimination performance (Perceptual Magnet Effect; Kuhl, 1991).

Insert Figure 2. around here

In order to minimize the ERP effects due to the pure physical properties of the vowels, and to diminish potential effects due to the order of presentation, averages for each contrast of interest were constructed to include both vowels of the pair in both positions (as standard and as deviant). For instance, in case of the V1-V4 contrast, the

standard condition included ERPs collected during the presentation of V1 (while V4 was the deviant) and V4 (while V1 was the deviant) as standards; naturally, the deviant condition included ERPs collected for V1 (when V4 served as standard) and V4 (when V1 was presented as standard). In this way, all participants contributed a minimum of 51 trials (out of 70) for each vowel (e.g., a minimum of 102 trials for each of the 12 main conditions reported below) contributed to the final averages and entered the statistical analyses.

ERP components were quantified by means of amplitude averages in three representative consecutive time windows (140–180, 210–260, and 440–580 ms), selected according to visual inspection of the grand average waveforms, and corresponding to time points associated with the N1, MMN, and the late negativity that emerged (see **Figure 3**). Using amplitude average as the dependent variable, separate Analyses of Variance (ANOVAs) were carried out for each time window (140-180 ms, 210-260ms, 440-450 ms) and vowel pair (V1-V2, V2-V3, V3-V4, V1-V4, V2-V4, and V1-V3) across 6 electrodes (F3, F4, Fz, C3, C4, and Cz) out of the 20, as the effects were most prominent at these sites and no additional components of interest emerged at the other sites. Each ANOVA included the within-subject factors *Condition* (standard, deviant) and *Electrode* (F3, F4, Fz, C3, C4, Cz) as well as the between-subject factor *Group,* with three levels based on the participants' language background (*monolingual French* (mF), *monolingual English* (mE), and *Simultaneous Bilingual* (SB)).

3. Results

As expected (based on **Figure 3**), ERP plots collapsed across all contrasts and all three groups were generally characterized by an N100-P200 complex in both the standard

and the deviant conditions, and the deviants elicited enhanced negativities in three distinct time intervals (140-180 ms, 210-260 ms, and 440-450 ms). Difference waves computed by conditions revealed the expected MMN component in most of the conditions, which began at around 210 ms and reached its peak around 230 ms. In addition, an early negativity with an onset around 140 ms and peak around 160 ms emerged in certain conditions, corresponding to the typical N1 auditory component. Finally, a late negativity peaking between 440 and 580 ms after stimulus onset is also visible in the difference waves of most conditions. We focus on the MMN results first, followed by the N1, and finally the results related to the late negativity.

● ●

Insert Figure 3 about here

● ●

3.1 MMN (210-260 ms)

In all Figures, the thin solid line represents the ERPs in response to the standard sounds, and the thick line shows the average ERPs in response to the deviant vowels. The dashed line represents the difference wave obtained by subtracting the standard condition from the deviant one revealing the MMN components that begin around 210 ms with an offset at about 260 ms.

Figure 4a illustrates the ERPs averaged across groups recorded in response to the contrast of V1 (French /u/) vs. V2 (English /u/), separately for each group. The difference wave (dashed line) reveals the emergence of an MMN between 210 and 260 ms, confirmed in an ANOVA, which yielded a significant main effect of condition ($F_{(1,42)}$=12.534; p=.001) with no significant group effect (F<1) or interaction (F<1).

Figure 4b displays ERPs computed for each language group for the V3 (control /y/) vs. V4 (French /y/) vowel contrast. The ANOVA revealed a main effect of condition ($F_{(1,42)}$ = 17.6; p = .001), as well as a group x condition interaction ($F_{(2,42)}$ = 3.651; p =.037). Follow-up analyses within groups showed that only the French group ($F_{(1,14)}$ = 6.42; p = .026), and the SB group ($F_{(1,14)}$=9.15; p =.01) exhibited a significant MMN response, whereas the English group did not (F<1).

. . . .
Figure 4. about here
. . . .

Figure 5a illustrates the V1-V4 contrast. The ANOVA revealed a main effect of condition ($F_{(1,42)}$ = 6.118; p = .017), with no group effect (F<1), nor any interaction (F<1), confirming that all three language groups showed a similar MMN between 210 and 260 ms. ERPs computed for the V2 (English /u/) vs. V3 (control /y/) contrast are illustrated in **Figure 5b**. The ANOVA revealed a main effect of condition ($F_{(1,42)}$ = 7.1; p = .01), as well as an interaction between group and condition ($F_{(2,42)}$= 2.02; p =.03). A clear MMN in the 210-260 ms time window was only recorded for the SB group. Follow-up analyses within the groups revealed a significant effect of condition only in the SB group ($F_{(1,14)}$= 6.5; p =.025), with no significant effects in either monolingual group (mE group, F<1, mF group, F<1).

. . . .
Figure 5. about here
. . . .

Results obtained in response to the other two contrasts (V1-V3, and V2-V4) that represent across-phonemic changes (with acoustic differences in both F1 and F2) are

illustrated in **Figure 6a and b.** For the V1-V3 vowel contrast (Figure 6a), the ANOVA revealed a main effect of condition ($F_{(1,42)}$ = 5.307; p = .027), as well as an interaction between group and condition ($F_{(2,42)}$= 3.8; p =.031). Follow-up analyses within the groups revealed a significant effect of condition in the SB group ($F_{(1,14)}$= 6.5; p =.021, and in the French group ($F_{(1,14)}$= 5.5; p =.031), with no significant differences in the English group (F<1). ERPs computed for the V2-V4 (control /y/) contrast are illustrated in **Figure 6b**. The ANOVA revealed a main effect of condition ($F_{(1,42)}$ = 4.082; p = .05), as well as an interaction between group and condition ($F_{(2,42)}$= 3.062; p =.036). Consistent with the ERP plots, follow-up analyses within the groups revealed a significant effect of condition only in the English group ($F_{(1,14)}$= 3.5; p =.049), and in the French group ($F_{(1,14)}$= 3; p =.05) with no significant differences in the SB group (F<1).

Figure 6. about here

In addition to the amplitude analyses, we examined peak latencies for all significant MMNs within each language group for each vowel pair. These values, which reflect the group average peak latencies found within the time window of the MMN, are displayed in each Figure. A one way ANOVA conducted on the MMN peak latencies showed no significant differences across language groups, with the exception of the V1 vs. V2 vowel pair, where a marginally significant main effect of language group was observed ($F_{(1,42)}$ = 3.309, p=.052). Post-hoc analyses contrasting the three groups showed that the MMN peak for the SB group (213 ms) occurred significantly earlier than those in the other two groups (SB vs. mE p=.05; SB vs. mF p=.05), whereas the two monolingual groups did not differ from each another (peak latencies in both groups: 247 ms).

3.2 N1 (140-180 ms)

Prior to the MMN effects, the ERP averages revealed that the deviant condition elicited a larger N1 component than the standard condition in response to the V1 vs. V4 and the V1 vs. V2 contrasts (clearly reflected in the difference waves). This occurred within the time window of 140 and 180 ms after stimulus onset, as highlighted by the first grey shaded area in the figures. In the case of the V1 (French /u/) vs. V4 (French /y/) contrast, a main effect of condition was observed ($F_{(1,42)}$=8.78; $p =$. 005) with no interaction with group, demonstrating that all three groups exhibited this N1 condition effect. Similarly, the main effect of condition approached significance ($F_{(1,42)} = 3.7$; $p =$ .0611) in the ANOVA for the V1 (French /u/) vs. V2 (English /u/) contrast, with no group interaction. No main effects or interactions were observed for the remaining four vowel pairs.

3.3 Late Negativity (440-580 ms)

Interestingly, discrimination effects were also signaled by a late negativity following the MMN components in 5 (out of 6) vowel contrasts. This late negativity effect is visible in the grand average ERPs highlighted by the third grey-shaded area, as well as in the corresponding difference waves, emerging at about 440 ms and lasting almost until 600 ms after stimulus onset.

The ANOVA for the V1 vs. V4 contrast (Figure 4) revealed that this late negativity was elicited across all groups, as reflected in a main effect of condition with no group interaction ($F_{(1,42)}= 10.122$; $p =$.003). A similar shared pattern emerged for the V1 vs. V2 contrast, as shown in Figure 4a (main effect of condition with no group interaction ($F_{(1,42)} = 13.412$; $p = $ .0001), and for the V3 vs. V4 comparisons, as shown in Figure 4b

(main effect of condition without group interaction: $F_{(1,42)}=10.1$; $p =.003$). In the case of the V1 vs. V3 contrast (Figure 6), the main effect of condition was also significant ($F_{(1,42)}= 11.03$; $p =.002$) with no group interaction, as well as in the V2 vs. V4 contrast ($F_{(1,42)}= 12.03$; $p =.001$).

4. Discussion

The overall goal of the current study was to investigate the long-term neural memory traces (as reflected by the MMN component) developed for native vowels when the overall language organization of the participants is based on acquiring one or two languages from birth. The experiment was designed to test: (1) whether the pre-attentive speech processing in SBs is tuned to the linguistic characteristics of the speech input from two languages, such that both within-language and cross-language phonetic differences are present in *a single phonetic space* in which access to sub-phonemic differences across the two languages is similar to that of monolingual peers; (2) whether SB listeners are sensitive to additional *sub-phonemic detail* in each language which might be irrelevant for monolinguals. The latter possibility is based on the assumption that the SBs' ability to adapt their perception to function optimally in different language contexts may require early, automatic access to a higher level of phonetic detail.

Given the overall ERP results, it is evident that simultaneous bilinguals and monolinguals (who are all native users of English and/or French) exhibit different neurophysiological patterns in response to native and cross-language vowels, in support of the holistic view of bilingualism (Grosjean, 1998). Discrimination of the vowels was signaled by different patterns of the MMN triggered in the time window of 210 to 260 ms

across the three language groups. Vowel discrimination was also apparent in the N1 effect (140-180 ms), and in a late negativity occurring between 440 and 580 ms after stimulus onset. The latter two ERPs showed no evidence of variation as a function of linguistic background.

Mismatch Negativity (MMN). Focusing on the MMN findings first, the within category phonemic contrasts (V1-V2 and V3-V4) produced results similar to previous investigations (i.e., Aaltonen et al., 1987; Csibra et al., 2000; Näätänen et al., 1997; Peltola et al., 2003. As expected, the English monolingual speakers did not show an MMN response in response to the V3-V4 contrast, as this part of the vowel space carries no relevant phonemic information in English (as a reminder, /y/ is not part of the English vowel inventory). The relatively small acoustic difference between these vowels (in terms of F1) appeared to be insufficient to elicit an MMN component in this group, (unlike the larger acoustic change between V1 and V4). However, the V3-V4 contrast (good vs. poor French /y/) did elicit an MMN response in the French and the SB groups, thus illustrating language-specific effects, as V3 and V4 are within-phoneme-category representatives for both the French and SB listeners. Since the MMN showed identical features in terms of latency and amplitude across the two groups, having acquired French as a monolingual or bilingual does not seem to lead to a different perceptual organization for this specific phonemic category, which has no close English 'neighbors'.

The V1-V2 stimulus pair represents an interesting phonetic contrast, as it contains the best exemplar of the English [u] (V2) and the best exemplar of the French [u] (V1), differentiated by the exact same acoustic-perceptual distance as the V3-V4 contrast. In response to this contrast, each group exhibited a reliable MMN. Similar MMNs across

the groups are expected because at least one of the vowels in this particular contrast is prototypical to the listeners. Interestingly, the SB listeners exhibited a faster peak latency compared to the other groups, pointing to an MMN response that is distinct from those of the monolingual listeners. This finding cannot be attributed to the SB group being generally faster than monolinguals in pre-attentive, automatic processing of speech sounds because this group difference was only present for the V1-V2 contrast, and not for the V3-V4 contrast, which represents a comparable acoustic-perceptual difference; the two sets of vowel contrasts only differ in terms of phonemic status.

Typically, the latency of MMN responses decreases when the stimuli pair crosses a phonemic boundary (Winkler et al., 2009). Therefore, this distinct neurophysiological response in terms of latency might suggest that SB listeners distinguish these vowels as different phonemic categories (English /u/ and French /u/) rather than as within-category variants (allophones). Several behavioral experiments have already illustrated that sequential bilinguals are able to perceptually separate acoustically overlapping cross-language categories when the ongoing linguistic context is highly controlled (e.g., Bohn and Flege, 1993; Elman et al., 1977; Garcia-Sierra et al., 2009). Even though such experiments have demonstrated that proficient sequential bilinguals are able to adjust their perception to the ongoing language context (L1 or L2), it has still remained unclear whether a bilingual's ability of perceptual adjustment was supported by separate phonemic categories across the languages, or whether bilinguals rely on allophonic information during the process.

The across-phonemic category contrasts that are defined by changes in F1 only (V1-V4, V2-V3) elicited different MMN patterns as a function of SB vs. monolingual

status only in the case of the V2-V3 vowel pair; the V1-V4 contrast elicited comparable MMNs across the groups. The largest acoustic difference extends between stimuli V1 and V4 (French /y/ vs French /u/) and corresponds to a clear cross-category (phonemic) difference in French and to a perceptible acoustic change (and/or a possible phoneme category change in English.) As shown by our previous behavioral investigations, the prototypical French [y] is categorized as a very poor exemplar of the English [i] by most English listeners. Therefore, the MMN observed in the English group might be due to categorical overlap between the French [y] and the English [i] vowels. Additionally, language differences are likely minimized by the large acoustic differences, given that ERPs recorded during the processing of contrasts with smaller acoustic differences (V1-V2, V3-V4, V2-V3, V1-3, and V2-V4) did reflect language-specific effects on the MMN, similar to previous investigations (i.e., Aaltonen et al., 1987; Csibra et al., 2000; Näätänen et al., 1997; Peltola et al., 2003;).

The finding that the monolingual English group exhibited an MMN only in response to either a large acoustic difference (V1 vs. V4) or to a native within-category change (V1 vs. V2), and the monolingual French and the SB group showed MMN responses to across- and within-category variations, implies that MMNs can be elicited by either a large acoustic difference or by contrasts where at least one of the vowels belongs to (or is a good category exemplar of) the native vowel system. Similar MMN effects have emerged in previous investigations in regard to consonant perception (i.e., Rivera-Gaxiola et al., 2000; Shafer et al., 2004).

The processing of the V2-V3 contrast revealed an important difference between SB and monolingual processing. The V2 and V3 vowels only differ in their F2 values,

sharing a common F1 value (300Hz); in comparison, the other two within-category contrasts (e.g., V1-V2 and V3-V4) differ in both F1 and F2 values. The SB listeners' MMN response to the V2-V3 pair suggests that the subtle difference between these two vowels is perceptually prominent for SB listeners only; thus SB listeners show sensitivity to a wider range of sub-phonemic differences (specifically, changes along the F2 axis of the vowel space) than their monolingual peers.

Interestingly, SB listeners in the current study appear to resolve and encode more detail in a part of the vowel space that is "crowded" when all native cross-language vowels are considered. Moreover, it is noteworthy that the SB listeners are resolving more detail in F2, given that the highest vowels in French (/i/, /u/ and /y/) and English (/i/, /u/) are differentiated *across* these languages primarily by their F2 values. Therefore, it is highly adaptive for SBs to be more sensitive to changes in F2 than their monolingual peers who encounter only a subset of these high vowels in a less crowded space. This finding is consistent with several previous studies in showing that SB listeners *differentiate*, rather than assimilate, highly similar phones across their languages (e.g., Guion, 2003; Sundara, Polka & Baum, 2006; Sundara and Polka, 2008). The present findings, similar to prior reports, also show that this perceptual adaptation in SB may involve resolving differences that are not phonemic in either language and appears to give rise to some unique, and highly efficient, perception (and production) patterns in comparison to either of their monolingual peers (e.g., Sundara et al, 2006; Sundara & Polka, 2008; Sundara, Polka & Molnar, 2008; Macleod & Stoel-Gammon, 2005).

The final pair of vowel contrasts (V1-V3, and V2-V4) that also represent across-vowel category changes (but are defined by changes in both F1 and F2) demonstrated

clear differences between the phonological organization of SBs and the monolingual listeners. The V1-V3 contrast represents the prototypical French [u] (V1) and the control vowel (V3; a very poor exemplar of the French vowel [y]). The English group showed no reliable MMN response, as expected, given that none of the vowels is prototypical in Canadian English; as opposed to the V1-V4 contrast, this acoustic difference was not sufficient to elicit an MMN irrespective of the non-phonemic status of the stimuli. The French and SB groups, however, showed comparable MMN patterns, also as expected, given the French phonemic status of V1.

The ERPs in response to the V2 and V4 contrast provided a unique insight into the phonological organization of monolinguals and SBs. This vowel pair (prototypical French [u] and prototypical English [u]) signals an across-phonemic category change, but also a switch between languages or linguistic contexts for the bilinguals. Similarly to the previous vowel pairs, when there was a prototypical vowel present in the contrast, the monolingual listeners exhibited an MMN pattern; the MMN responses of SBs were attenuated, however. Similar patterns in monolinguals have been demonstrated before; monolingual German listeners' processing of a non-native (Polish) deviant consonant was weaker in the presence of a linguistically relevant native contrast, and their responses to a native deviant consonant was increased by the presence of non-native sounds (Lipski and Mathiak, 2008). In the present study, it appears that the presence of either a prototypical French or English vowel as standard (played 70% of the time) defined a linguistic context in the oddball paradigm. Listeners exhibited MMNs when the deviants represented within-language and within-phonemic category changes (e.g., V1-V4, V1-V2, and V3-V4), although this particular contrast represented a clear change between two linguistic

modalities of SBs. Accordingly, it seems that when V2 or V4 served as standards, SBs showed MMN responses to within-linguistic context variation but showed an attenuated MMN for a vowel that clearly belonged to another linguistic modality. In this case, the Late Negativity signaled auditory discrimination only.

Accordingly, it is possible that the phonemic status of vowels in a multiple oddball paradigm defines an internal linguistic context (independently of the presence or absence of an external linguistic context), and affects the pre-attentive auditory processing of bilingual listeners. This implies that SBs do not rely on a completely shared phonological space across their languages (contra Peltola et al., 2012), and that language context affects bilinguals' vowel processing, as has been demonstrated in several behavioral paradigms (e.g., Bohn and Flege, 1993; Elman et al., 1977; Garcia-Sierra et al., 2009).

Late Negativity in the ERPs. In addition to the MMN findings, interestingly, discrimination abilities across all the vowels and language groups were reflected in a late negativity as well (except for V2-V3). The late negativity was reliably elicited in the monolingual speakers even when no MMN was present. The overall pattern of the late negativity suggests that it might be related to, but not dependent upon, the MMN component, as it can appear in the absence of the MMN as well; and it appears to signal a process related to discrimination, as it is elicited by sounds with deviant status in the oddball paradigm.

Similar late negativities have been reported in the EEG literature (e.g., Alho et al, 1992; Čeponienė et al., 1998; Korpilahti, 1997; Trejo et al, 1995); however, the exact role of this component remains unclear. Studies investigating infants' and children's auditory

perception often record a late negativity (termed late discriminative negativity) peaking between 400 and 500 ms (for review, see Cheour et al, 2001), that has been shown to be greater in response to speech sounds than to tones of equal complexity (e.g., Korpilahti, 1997). As the late discriminative negativity has been linked to developmental changes - given that it is obtained most reliably in children and its amplitude shows a decrease with age and seems to completely disappear by adulthood (Cheour et al, 2001) - it may not be directly related to the late negativity recorded in the current study. However, it is important to note that Čeponienė et al. (1998) demonstrated that the second negativity in children is more prominent when the ISI is longer; consequently, the duration of the ISI in the present investigation might have influenced the emergence of a second negativity in adults as well.

Only a few adult studies have reported a late negativity in a non-attended paradigm (Alho et al, 1992; Trejo et al, 1995), and it is somewhat unclear whether these negativities were related to the standard stimuli preceding the deviants (a "spill-over" processing of the previous sound) or whether they reflected a discriminatory process associated with the deviant sound, since these studies relied on a short ISI (< 400 ms). In the current study, due to the elongated ISI between standards and deviants, the late negativity cannot be related to the processing of the preceding sounds (as the late negativity occurred before the presentation of the following stimuli), and it is more likely associated with change detection in the auditory input. In addition, given that the SB group exhibited a somewhat different pattern for the late negativity as compared to monolinguals, this late component might reflect the effects of language experience as well.

N1 component. In one contrast (V1-V4), a significant N1 difference was recorded. The auditory N1 component has been shown to be primarily sensitive to the physical and temporal characteristics of the auditory stimuli; however, certain subcomponents of the N1 have been shown to be related to condition-dependent effects (for review, see Näätänen and Picton, 1987). Recall that we relied on an averaging method that ruled out any negativities in the final grand averages due to the physical characteristics of the stimuli; therefore the N1 effect in our ERPs more likely reflects effects related to the characteristics of the given conditions.

Long ISIs have also been shown to modify the N1 component; for instance, Czigler and colleagues (1992) reported enhanced N1 amplitudes to changes in tones when the relatively long ISIs of 800, 2400 and 7200 ms were employed in different experimental blocks. Additionally, Rivera-Gaxiola et al. (2000) recorded N1 effects in response to certain contrasts when using a long ISI of 1500 ms to illustrate cross-linguistic effects on consonant perception during a passive auditory oddball paradigm. In the current study, the N1 was most prominent in response to the largest acoustic change in the stimulus set, thus it is a possibility that the N1 effect is related to acoustic changes and not to phonemic information.

Conclusion. The overall MMN findings demonstrate three unique characteristics of the English-French SB perceptual system. First, acoustically similar cross-language vowel categories present in the SB speech input appear to be stored as separate categories (rather than within-category variants) in the form of long-term phonemic memory traces that are retrieved automatically during the neural path of speech processing at the pre-attentive level. Similar findings emerged from behavioral studies investigating both the

production and perception of acoustically similar speech sounds (i.e., Guion, 2003; Sundara et al., 2006; Sundara et al., 2008), supporting the idea that bilingual speakers, instead of assimilating overlapping cross-language categories, rather tend to maintain separate categories (dissimilation). This has been primarily shown in behavioral studies when the ongoing language context was optimized to display such effects (i.e., Elman et al., 1977; Bohn and Flege, 1993; Garcia-Sierra et al., 2009), unlike in the current study where external language context and other task effects were minimized.

Second, SB listeners appear to be more attuned to sub-phonemic changes as compared to monolinguals, most obvious in the V2-V3 contrast. Because this contrast contains a prototypical English vowel ([u]), it was expected to elicit a significant MMN response at least in the English and SB group. However, only the SB group revealed a MMN. It appears that relatively small acoustic changes in one dimension only (F2) were not perceptually salient enough for the monolingual English listeners to elicit a statistically reliable MMN when the stimuli were presented with a long ISI. Changes in the F2 value were more relevant to the SBs as compared to the monolinguals. This can be explained by the specific organization of the English-French vowel space: F2 is the most relevant cue for differentiating the English and French vowels in this particular part of the vowel space. Therefore, SB listeners utilize the most relevant phonetic information across the two languages at the very early stages of speech perception.

Third, when the standard vowels provided a clear cue for linguistic contexts, the SB listeners' pre-attentive processing (at the level of the MMN) appeared to be less sensitive in response to vowels that were not part of the specific context. Therefore, it appears that linguistic context (without the presence of an external context) specified by

the standard sounds of the oddball paradigm affects SBs' vowel discrimination, and the passive auditory oddball-paradigm itself can create an internal linguistic context. Also, it appears that SBs are sensitive to context effects at an early stage of processing as reflected by the MMN, but only when non-ambiguous linguistic cues are provided (e.g., V2 vs. V4), but not when the linguistic cues are ambiguous (e.g., V1 vs. V2).

In sum, results indicate that the pre-attentive process involved in SB vowel perception is shaped to detect the phonetically most salient information that is specific to each of the languages, as demonstrated by the MMN responses to the V1-V2 and V2-V3 contrasts. SB listeners of English and French exhibit a speech perception pattern that is distinct from both monolingual listener groups even during the earliest levels of speech processing, as the SB pre-attentive system is tuned to access sub-phonemic detail with respect to both of their input languages, including detail that is not readily accessed by either of their monolingual peers. This automatic access to fine phonetic detail may be essential in supporting the SBs' ability to make rapid, effortless shifts in perception across different communication/linguistic contexts. Moreover, bilinguals appear to function in a hybrid phonological space: when a clear linguistic context is present, the context-appropriate information is the most salient; when no clear context is established, bilinguals show heightened sensitivity to cues that are salient for discriminating acoustically similar cross-language categories.

Acknowledgements

5. References

Alho, K., Woods, D., Algazi, A., & Näätänen, R. (1992). Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central spacestar open. *Electroencephalography and Clinical Neurophysiology*, 5, 356-368.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 171-206). Baltimore: York Press.

Bohn, O.-S., & Flege, E. J. (1993). Perceptual switching in Spanish/English bilinguals. *Journal of Phonetics*, 21, 267-290.

Bosch, L., & Sebastián-Gallés, N. (2003). Evidence of early language discrimination abilities in infants from bilingual environments. *Language and Speech*, 46, 217-243.

Bongaerts, T., Mennen, S., Slik, F van der (2000). Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced late learners of Dutch as a second language. *Studia Linguistica* 54, 1467-9582.

Burns, T. C., Werker, J. F., & McVie, K. (2003). Development of phonetic categories in infants raised in bilingual and monolingual environments. Paper presented at the Boston University Conference on Language Development, Boston.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of Acoustical Society of America*, 54(2), 421-428.

Carney, A., Widin, G., & Viemeister, N. (1977). Noncategorical perception of stop consonants differing in VOT. *Journal of Acoustical Society of America*, 62, 961-970.

Čeponienė, R., Cheour, M., & Näätänen, R. (1998). Interstimulus interval and auditory event-related potentials in children: evidence for multiple generators. *Electroencephalography and Clinical Neurophysiology*, 108, 345-354.

Cheour, M., Korpilahti, P., & Martynova, O. (2001). Mismatch negativity (MMN) and late discriminative negativity (LDN) in investigating speech perception and learning in children and infants. *Audiology and Neurotology*, 6, 2-11.

Czigler, I., Csibra, G., & Csontos, A. (1992). Age and inter-stimulus interval effects on event-related potentials to frequent and infrequent auditory stimuli. *Biological Psychology*, 33, 195-206.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport*, 8(4), 919-924.

Flege, E. J. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 233-272). Baltimore: York Press.

Flege, J., Yeni-Konshian, G. H., & Liu, S. (1999). Age constraints on Second-Language Acquisition. *Journal of Memory and Language*, 41, 78-104.

Flege, E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, 40(4), 467-491.

Garcia-Sierra, A., Ramirez-Esparza, N., Silva-Pereyra, J., Siard, J., Champlin, C. A. (2012). Assessing the double phonemic representation in bilingual speakers of Spanish and English: An electrophysiological study. *Brain and Language*, 121, 194-205.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1, 131-149.

Guion, S. G. (2003). The vowel systems of Quichua-Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica*, 60, 98-128.

Harris, R., & Nelson, M. (1992). Bilingualism: Not the Exception Any More. *Advances in Psychology*, 83, 3-14.

Korpilahti, P., Salmela, S., Lang, H., Porn, B., & Crause, C. (1997). Event-related potentials elicited by complex tones, words and pseudo-words in normal and language impaired children. *Electroencephalography and Clinical Neurophysiology*, 103, 64.

Kuhl, P. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not1. *Perception and Psychophysics*, 50, 93-107.

Lipski, S. C., & Mathiak, K. (2008). Auditory mismatch negativity for speech sound contrasts is modulated by language context. *NeuroReport*, *19*(10), 1079-1083.

MacKay, I. R. A., Meador, D., & Flege, E. (2001). The Identification of English Consonants by Native Speakers of Italian. *Phonetica*, 58, 103-125.

MacLeod, A. A., & Stoel-Gammon, C. (2005). Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders*, 3, 118-127.

Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50, 940-967.

Molnar, M, Polka, L., Baum, S., & Steinhauer, K. (2010). Vowel perception: How simulatnous bilinguals do it. Neurobilingualism Conference. Donostia, Spain. September 30 – October 1.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Livonen, A., et al. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385, 432-434.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neuropsychology*, 118, 2544-2590.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic respond o sound: a review and an analysis of the component structure. *Psychophysiology*, 24, 375-425.

Pallier, C., Bosch, L., & Sebastián-Gallés , N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, 64, B9-B17.

Peltola, M. S., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., & Näätänen, R. (2003). Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. *Neuroscience letters*, 352, 25-28.

Peltola, M. S., Tamminen, H., Toivonen, H., Kujala, T., Näätänen, R. (2012). Different kinds of bilinguals – Different kinds of brains: The neural organisation of two languages in one brain. *Brain and Language*, 121, 261-266.

Perani, D., Paulesu, E., Sebastián-Gallés , N., Dupoux, E., Dehaene, S., Bettinardi, V., et al. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain*, 121, 1841-1852.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, 13, 253-260.

Piske, T., Flege, E., MacKay, I. R. A., & Meador, D. (2002). The Production of English Vowels by Fluent Early and Late Italian-English Bilinguals. *Phonetica*, 59, 49-71.

Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, 41, 221-231.

Polka, L. & Bohn, O-S. (2011) Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, 39, 467-478.

Polka, L., Molnar, M., Baum, S., Menard, L., & Steinhauer, K. (2009). Asymmetries in the MMN response to vowels by French, English, and bilingual adults: evidence for a language-universal bias. Paper presented at the Acoustical Society of America Meeting. Portland, Oregon, USA.

Rivera-Gaxiola, M., Csibra, G., Johnson, M., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behavioral Brain Research*, 111, 13-23.

Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology*, 62, 437-448.

Sharma, A., & Dorman, M. (2000). Neurophysiologic correlates of cross-language phonetic perception. *Journal of Acoustical Society of America*, 107(5), 2697-2703.

Sebastián-Gallés , N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: is the distributional account all there is to it? *Developmental Science*, 1467-7687.

Sebastián-Gallés, N., Echeverría, S., Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, 52, 240-255.

Sundara, M. (2005). Acoustic-phonetics of coronal stops: A cross-language study of Canadian English and Canadian French, Journal of the Acoustical Society of America, 118, 1026-1037.

Sundara, M., & Polka, L. (2008). Discrimination of coronal stops by bilingual adults: The timing and nature of language interaction. *Cognition*, 106, 234-258.

Sundara, M., Polka, L., & Baum, S. (2006). Production coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition*, 9, 97-114.

Sundara,M., Polka, L., & Genesee, F. (2006). Language experience facilitates discrimination of /d*D*/ in monolingual and bilingual acquisition of English. *Cognition, 100*, 369–388.

Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, 108, 232-242.

Steinhauer, K., Drury, J. (2011). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, 120, 135-162.

Trejo, L., Ryan-Jones, D., & Kramer, A. (1995). Attentional modulation of the mismatch negativity elicited by frequency differences between binaurally presented tone bursts. *Psychophysiology*, 32, 319-328.

Werker, J., & Logan, J. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics*, 37, 35-44.

Winkler, I., Kujala, T., Paavo, A., & Näätänen, R. (2003). Language context and phonetic perception. *Cognitive Brain Research*, 17, 833-844.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., et al. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, 36, 638-642.

Yamada, T., Yamada, R., & Strange, W. (1995). Perception of English vowels and consonants by Japanese learners of English. Paper presented at the Proceeding of the Acoustical Society of Japan.