Perceiving non-native vowel contrasts: ERP evidence of the effect of experience

Cheryl Frenck-Mestre, Pauline Peri, Christine Meunier and Robert Espesser
Laboratoire Parole et Langage, CNRS, Université de Provence

ABSTRACT

In two ERP experiments, we used a 3-stimulus oddball paradigm to examine the perception of American-English vowel contrasts, by native speakers, French monolinguals and late French-English bilinguals. In the first experiment, stimuli were standard /ɛ/ (75%), target /ae/ (10%) and oddball /ɪ/ (15%). In the second, the attentional demands were inversed: standard /ɛ/ (75%), target /ɪ/ (10%) and oddball /æ/ (15%). In both experiments and in all 3 groups, early acoustic discrimination of all vowels was shown by variations in the N100 response. Subsequent phonemic categorization as revealed by the P300 response differed, however, across experiments and groups. Bilinguals showed a P300 response to oddball /ɪ/ akin to English speakers, though reduced in magnitude and distribution, whereas French monolinguals did not. When /ɪ/ became the target and /æ/ the deviant, native English and French showed an inversion of the P300 to these vowels; bilinguals could not disengage attention and showed equal P300 responses to both. Our results indicate that while late French-English bilinguals establish new L2 vowel categories these categories are not as stable as L1 categories. Late learners' ability to discriminate and to selectively attend to a given L2 category is intermediate between that of native speakers and native speakers of their own L1.

KEYWORDS: ERPs, P300, late bilinguals, non-native contrasts

1. Introduction

One of human being's irrepressible activities is that of categorisation, the act of consciously sorting the observable events in the external environment into recognizable entities that can then be further processed, stored in memory as identifiable traces and recalled from memory as needed. The perception of the sounds of the world's languages is no different. From infancy, we learn to differentiate auditory events and to categorize them as either exemplars of a common category or as members of different categories (Dehaene-Lambertz and Baillet, 1998; Kuhl et al., 1992; Werker and Tees, 1984; Werker, 1994). The results of numerous studies, obtained with a wide variety of techniques, suggest moreover that one's early experience with auditory input has considerable impact on the ability later in life to create new categories, otherwise stated to reorganize the existing system to accommodate novel sounds as opposed to assimilating them to established categories (Iverson and Kuhl, 1995; Dehaene-Lambertz, 1997) although, as outlined below, various factors over and above early-life experience play a role in how easily new perceptual categories can be created. The present study used event-related cortical potentials to determine the extent to which adult learners create new perceptual categories when they learn a foreign language, how native-like their categories are and the role of experience in this ability.

There is ample evidence that non-native contrasts can be acquired after infancy; the extent of success depends nonetheless on numerous factors. Early exposure, during childhood, to a second language does not in fact guarantee the formation of native-like phonemic categories; however, the likelihood of such is increased if the L1 is not maintained (Flege and McKay, 2004; Tsukada et al., 2005). L2 contrasts that form a “single category” in the L1 may not become native-like even for early learners (Højen and Flege, 2006). Moreover, the acquisition of L2 vowel contrasts has often been found to be conditioned by the phonological features present in the native language of learners (Brown, 1998; McAllister et al., 2002). These results, which support the idea that the native language plays an important role and enables one to predict the relative difficulty of acquisition of a given L2 contrast, are in line with two predominant models of adult learning and perception of non-native contrasts. Both the Speech Learning Model, or SLM (Flege 1995), which focuses on second language acquisition, and the Perception-Assimilation Model, or PAM (Best, 1994), which is aimed at the study of perceptual capacities in monolinguals (but see Best & Tyler, 2007, for an extension to L2 learning), provide a rich theoretical framework of non-native phonemic perception. Note that neither model restricts the acquisition of non-native contrasts to the simple presence vs. absence of phonological contrasts within the L1. Indeed, a considerable body of work has shown that the capacity to perceive L2 vowel contrasts is influenced by phonetic properties of the speech signal in and of itself such that L2 contrasts can be perceived in certain phonetic contexts/positions but not others (Strange, 2007; Levy and Strange, 2008; Trofimovich et al., 2001). Moreover, recent psycholinguistic studies of L2 perception have shown that meta-linguistic factors can also play a role in the discrimination of L2 minimal pairs (Weber and Cutler, 2004).

The present study presents further evidence that adult L2 learners can reshape their perceptual space to accommodate L2-specific vowel contrasts but that, in line with the results of the vast majority of studies, “native-like” performance is not observed. Whereas the majority of studies of L2 (and non-native) phoneme perception have used behavioural techniques to address this question, we chose to record event-related potentials (ERPs) to examine the perceptual capacities of our participants. ERPs provide not only a millisecond precise measurement of processing but, depending upon the task, can allow one to disentangle automatic detection from attentional processing. To date, only a handful of ERP studies have examined phonemic processing specifically for L2 acquisition. These studies have used an oddball paradigm under conditions of non-attentional processing and examined the mismatch negativity response (MMN) to L2 contrasts. The results of these studies are mixed. Whereas Winkler et al. (1999) found that adult late L2 learners who had been immersed for several years in their L2 perceived non-native contrasts (in Finnish) as well as native speakers, this result was not replicated in a population of advanced adult L2 learners (of English) who were not immersed (Peltola et al., 2003). The results obtained for Finnish children in either bilingual or full immersion programs revealed very rapid onset of neural changes in response to L2 (French) vowel contrasts (Cheour et al., 2002; Peltola et al., 2005). Again, however, subsequent work did not confirm these findings when the L2 was English (Peltola et al., 2007), although the pattern of results of the bilinguals was puzzling. Nonetheless, caution is warranted before concluding that young L2 learners achieved native-like capacity to perceive their L2 as no L1 control group was included in these studies.

In this paper, we report the results of two ERP investigations of non-native vowel processing, using an active oddball paradigm. Rather than have our participants passively attend to auditory stimuli, we required them to mentally count a target vowel (10% of trials) presented against a background "standard" (75% of trials) and a deviant (15%) vowel. This allowed us to examine the effects of attention, by varying the status of a given vowel, i.e. whether target or deviant, across experiments. To address the question of just how "native-like" L2 learners' perception can become, we compared 3 groups: control native English speakers, control French speakers who had only had classroom learning of English, and French-English late bilinguals.

2. Method

2.1. Participants

Twenty-four right-handed adults aged 18 to 24 participated in the study. There were 8 English speakers living in France (mean of 6 months), 8 native French "monolinguals" who had learned English as a foreign language throughout secondary school, and 8 late French-English bilinguals. All late bilinguals were native French speakers raised by native French parents, who had learned English in secondary school starting at age 11, were studying to become English instructors, had lived at least 1 year in an English speaking country and rated themselves as fluent in English for oral and written comprehension and expression.

2.2. Stimuli

The stimuli consisted of the American-English vowels /ɛ/, /ɪ/ and /æ/, extracted from a lexical context (/h/_/d/) produced by a trained female native speaker. Twenty-five utterances of each lexical item were recorded in a sound attenuated room and digitized at 32 kHz/32 bits. Stimuli were created by extracting 60 ms from either side of the steady-state of the vowel and adding 10 ms contours to the onset and offset. Seven tokens were selected for each vowel category. F0 was held constant across stimuli. F1, F2 and F3 values were automatically extracted. The mean formant values were: /ɛ/ (F1 = 625 Hz, F2 = 2120 Hz, F3 = 2845 Hz), /ɪ/ (F1 = 480 Hz, F2 = 2295 Hz, F3 = 2953 Hz), and /æ/ (F1 = 930 Hz, F2 = 1857 Hz, F3 = 2698 Hz). The Euclidean distance calculated between mean Bark values for F1 and F2 was larger between /æ/ and /E/ (2.32) than between /ɪ/ and /ɛ/ (1.32).

2.3. Design and Procedure

Stimuli were delivered binaurally via headphones. A total of 625 stimuli were presented in a fixed random order, with a probability of occurrence of .75, .10 and .15 for standard /ɛ/, target and deviant stimuli respectively. Stimulus duration was 140 ms and SOA 1200 ms. In Experiment 1, the target was /æ/ and the deviant /ɪ/; in Experiment 2 the target was /ɪ/ and the deviant /æ/. Participants mentally counted the target stimuli. They were not informed of the presence of the deviant.

2.4. Electroencephalographic recording

EEG activity was recorded continuously at pre-frontal (Fp1, Fp2), frontal (F3, F4), occipital (O1, O2) and midline (Fz, Cz, Pz) electrode sites, referenced to the left mastoid. The EEG was amplified with a bandpass of 0.1 – 40 Hz (3dB cutoff) and digitized on-line at 200 Hz. EEGs were filtered offline below 15 Hz. Epochs began 100 ms prior to stimulus onset and continued 1100 ms thereafter. Average ERPs were formed off-line from trials free of muscular and/or ocular artifact.

2.5. Data Analysis

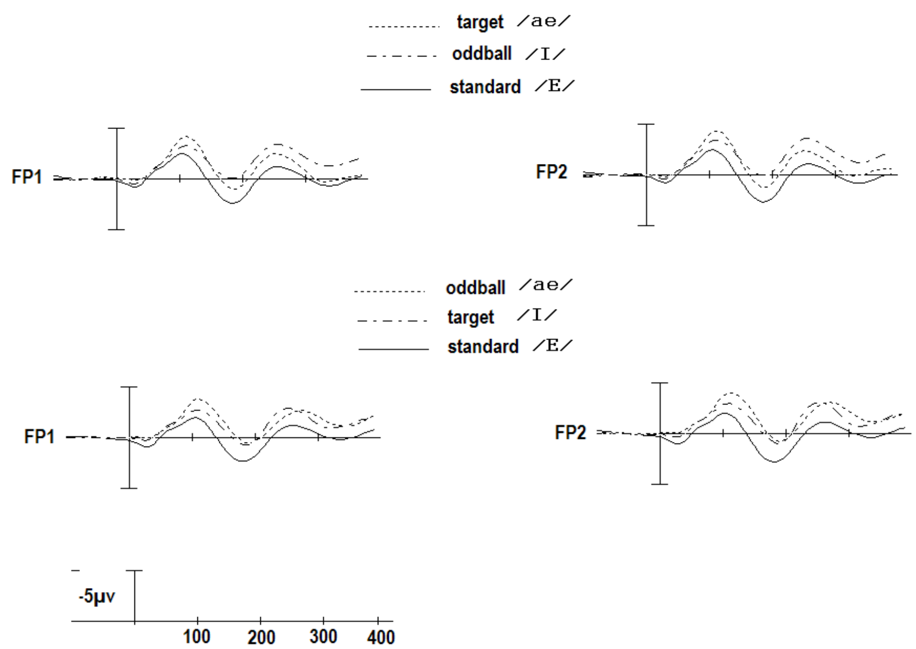
The ERP data were quantified by calculating the peak and/or mean amplitudes and latencies post-stimulus onset and relative to a 100 ms pre-stimulus baseline, for N100 (80–150 ms), for N200 (180–280 ms) and for P300 (280–500 ms) for each participant. Independent ANOVAs (Greenhouse-Geisser corrected) were performed on amplitude and latencies.

3. Results

3.1. N100

The grand average waveforms elicited by /ɛ/, /æ/ and /ɪ/ are presented for N100 as a function of experiment in Fig. 1. In both experiments, the effect of vowel was significant (Exp 1: F (2, 42) = 12.54 p<.001; Exp 1: F (2, 42) = 11.54 p<.01); N100 peak amplitude was largest for /æ/, smallest for the standard /ɛ/ and intermediate for /ɪ/, as confirmed by pairwise comparisons (p<.01 or better, Bonferroni). No interactions obtained with group (F<1) or experiment (F<1). The differences between the three vowels are in line with the Euclidean distance between them, and are independent of the role played, whether target or deviant.

Figure 1. Grand average wave forms as a function of vowel and experiment, for all participant groups.



3.2. P300

The grand average waveforms elicited by /ɛ/, /æ/ and /ɪ/ are presented for each of the three participant groups in the P300 window, as a function of vowel and experiment, in Fig. 2–4. In Experiment 1, the target /æ/ elicited a large positive deflection compared to the standard /ɛ/ in all groups. The deviant /ɪ/, however, only elicited a P300 response for native English and French-English bilingual participants. In Experiment 2, where participants were now requested to mentally count the vowel /ɪ/, both the target /ɪ/ and deviant /æ/ produced a larger mean P300 response than standard /ɛ/ in all groups. Importantly, the P300 response varied as a function of the role of the vowel across experiments, as confirmed by ANOVAs.

3.3. Experiment 1

Mean P300 amplitude varied significantly at midline as a function of Vowel ($F(2, 42) = 50.48$, $p<.0001$), which was modified by the higher order interaction involving Vowel, Electrode and Group ($F(8, 84) = 4.05$, $p<.001$). Post hoc comparisons revealed that in all groups, the target /æ/ produced a larger P300 response than standard /ɛ/ at all midline sites. For the deviant /ɪ/, in both the French-English bilingual group and the native English group this vowel produced a larger P3 amplitude than standard /ɛ/ at Cz and Pz. In the French monolingual group the deviant /ɪ/ did not differ from the standard /E/.

3.4. Experiment 2

Mean P300 amplitude varied significantly at midline as a function of Vowel ($F(2, 42) = 15.83$, $p<.0001$), which was modified by the higher order interaction involving Vowel, Electrode and Group ($F(8, 84) = 2.44$, $p<.05$). Post hoc comparisons revealed that in the native English group, mean P300 amplitude was larger for both target /ɪ/ and deviant /æ/ than standard /ɛ/, but that the effect was larger for the target /ɪ/. In the French monolingual group, target /ɪ/ produced a larger P300 than standard /ɛ/ whereas the deviant /æ/ tended to (p.07). In the French-English bilingual group, both target /ɪ/ and deviant /æ/ produced larger P300 responses than the standard /ɛ/, and did not differ from each other.

3.5. Experiment 1 vs. Experiment 2

To assess the effect of task demands, i.e. of the allocation of attentional resources to perceive the different vowel categories, an ANOVA was performed on the effect sizes for the target and deviant vowels, i.e. the ERP response produced in response to each of these stimuli in comparison to the standard vowel /ɛ/, in the P300 time window across experiments. For native English speakers, the effect size of /æ/ and /ɪ/ was inversed across experiments in accordance with the task; for native French, the same inversion was found. For French-English late bilinguals, however, whereas the effect size of /ɪ/ increased Experiment 2 that of /æ/ did not decrease significantly.

Figure 2. P300 response as a function of vowel and experiment, for native English speakers.

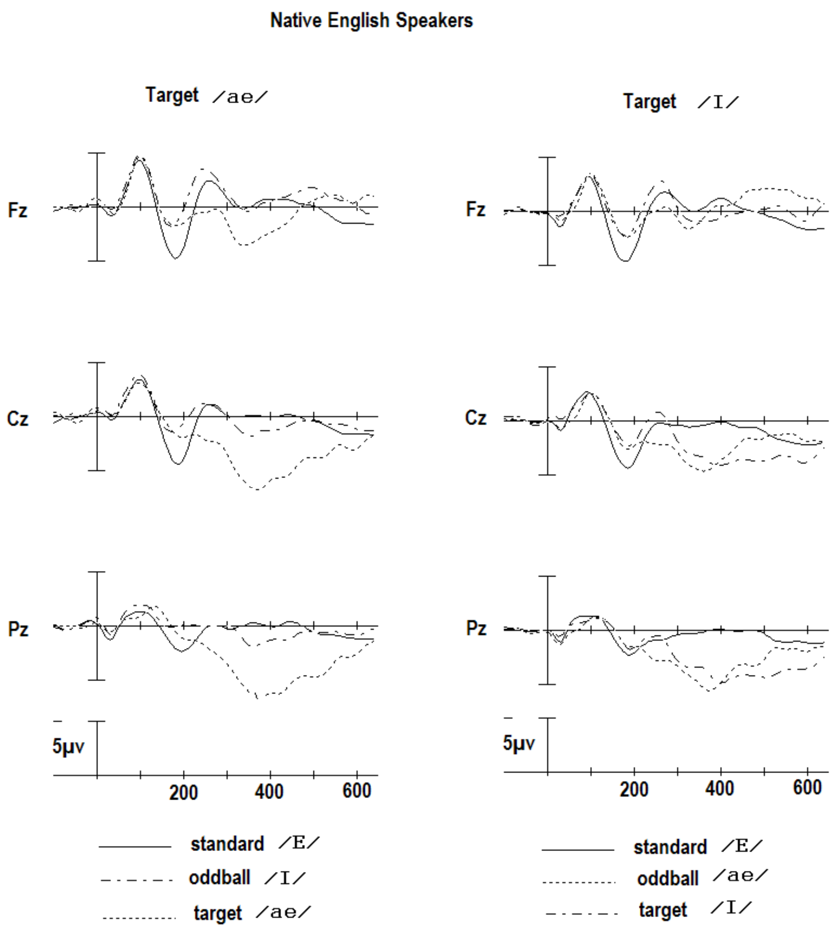


Figure 3. P300 response as a function of vowel and experiment, for French-English bilinguals.

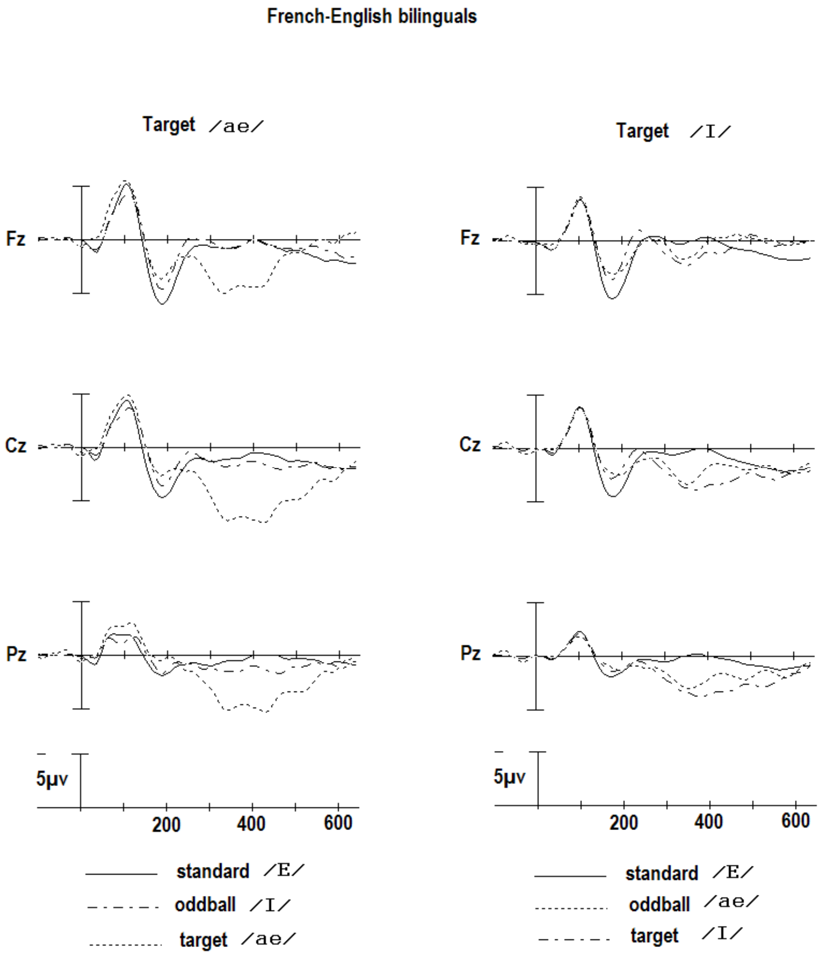
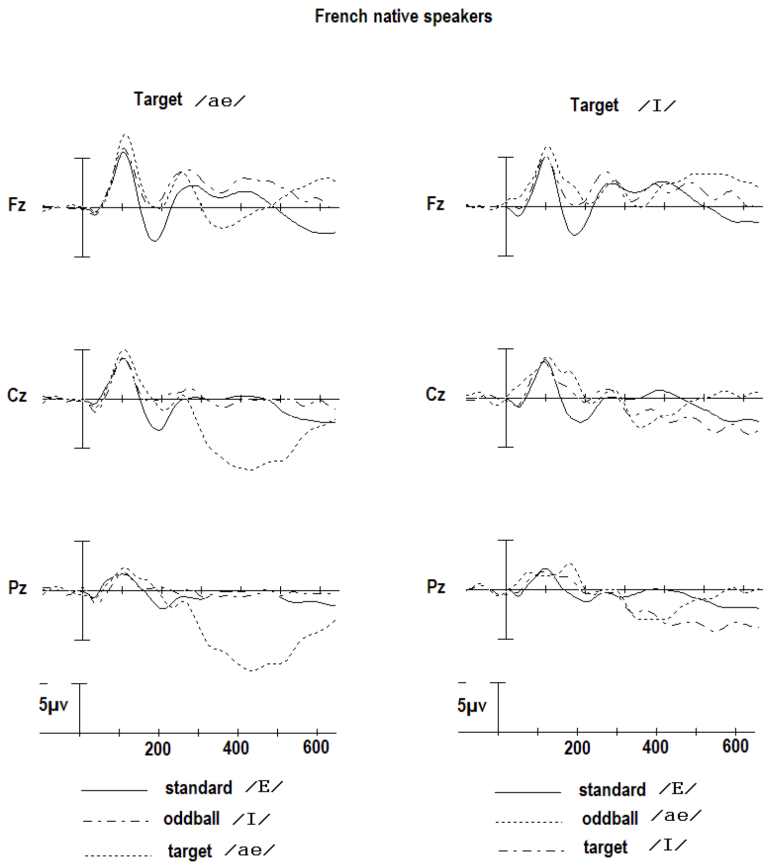


Figure 4. P300 response as a function of vowel and experiment, for French monolinguals.



4. Discussion

The results from the two experiments reported here provide important insight into the perception of vowel contrasts as a function of numerous factors: the status of these contrasts (i.e. whether native or not part of the listener's native repertoire), the status of the listener (i.e. whether performing as a monolingual or a bilingual), the type of processing (i.e. whether acoustic or phonetic), and task demands. Overall, the pattern of results show that our late French-English bilingual participants were in an intermediary position between native English speakers and French native speakers who have been exposed to English through secondary education but who are not proficient in this language.

For all participants, during the first 100 ms following presentation of the three American-English vowels we found a graded ERP response, which closely followed the acoustic differences between these vowels. This graded N100 response was impervious not only to the listeners' native language repertoire but also to task demands, such that it can be considered a likely "acoustic" response rather than related to categorical perception. This result replicates previous results for the same contrasts (Frenck-Mestre et al., 2005). Following this initial response, however, the three participant groups showed quite different patterns as revealed by the P300 response. Late French-English bilinguals showed a categorical response to the English vowel /ɪ/ against the background /ɛ/ even when attention was devoted to the target vowel /æ/. In this sense, they mimicked the pattern of native English speakers and differed from French monolinguals, who did not consciously discriminate /ɪ/ from /ɛ/ under these conditions. The French-English bilinguals differed from both the native English speakers and French monolinguals however, when task demands were switched. For the two groups working under "monolingual" conditions, the P300 response to the non-target vowel was significantly smaller than that to the target. For the bilingual group, even when attention was switched both American English vowels, /æ/ and /ɪ/ the vowel /ɪ/ elicited a reliable P300 response, with no reliable decrease in response to the non-target vowel (/æ/) across experiments, as though they were unable to reliably disengage attention. This pattern suggests that for the bilingual group, consciously sorting out the three vowel categories necessitated attention to all vowels in contrast to the "monolinguals" who were able to devote their attention exclusively to one category.

It is important to note that the comparison of the two French native speaker groups is one between "foreign language acquisition", i.e. for the group of French speakers who had predominantly been exposed to English in a classroom setting, and "second language acquisition", i.e. for the late bilingual group, who had all lived for at least 12 months immersed in the English language in addition to extensive classroom learning. Our results support the findings of ERP studies which have examined both pre-attentive and conscious processing of second-language and non-native contrasts in showing that simple exposure to these contrasts is not sufficient to produce a reliable automatic response but that immersion appears to be critical (Frenck-Mestre et al., 2005; Peltola et al., 2003). Nonetheless, even for our advanced bilingual group we did not find a pattern of vowel categorization that entirely overlapped with that of native speakers. In our first experiment, although the bilinguals consciously categorized the deviant, non-attended vowel contrast just as native speakers did, their electrophysiological response was smaller in both amplitude and distribution. In the second experiment, while the native speakers were able to selectively attend to just one of the English vowels, the group of late bilinguals did not demonstrate this capacity. These results suggest, as has been forwarded by Flege and McKay (2004) that it may be too much to expect late learners to achieve native-like performance but that they can indeed achieve very high levels of proficiency.

REFERENCES

Best, C.T. 1994. The emergence of native-language phonological influences in infants: a perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: the transition from speech sounds to spoken words*. Cambridge: MIT Press. 167–224.

Best, C. T., Tyler, M.D. 2007. Nonnative and second language speech perception. In O.S. Bohn and M.J. Munro (eds.) Language experience in second language speech learning. John Benjamins, Amsterdam, 13–34.

Brown, C. 1998. The role of the L1 grammar in the L2 acquisition of segmental structure. Second Language Research, 14, 136–193.

Cheour, M., Shestakova, A., Alku, P., Ceponiene, R., Näätänen, R. 2002. Mismatch negativity shows that 3–6-year-old children can learn to discriminate non-native sounds within two months. *Neuroscience Letters*, 325, 187–190.

Dehaene-Lambertz, G., Baillet, S. 1998. A phonological representation in the infant brain. *NeuroReport*, 9, 1885–1888.

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

Flege, J.E. 1995. Second-language Speech Learning: Theory, Findings, and Problems. In W. Strange (Ed) *Speech Perception and Linguistic Experience: Issues in Cross-language research*. Timonium, MD: York Press, 229–273.

Flege, J.E., McKay, I. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26, 1–34.

Frenck-Mestre, C., Meunier, C., Espesser, R., Daffner, K., Holcomb, P. 2005. Perceiving nonnative vowels: The effect of context on perception as evidenced by event-related brain potentials. *J. of Speech, Language, and Hearing Research*, 48, 1–15.

Højen, A., Flege, J. 2006. Early learners' discrimination of second-language (L2) vowels. *J. of the Acoustical Society of America*. 119, 3072–3084.

Iverson, P., & Kuhl, P. K. 1995. Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *J. of the Acoustical Society of America*, 97, 553–562.

Kuhl, P. K., Williams, K.A., Lacerda, F., Stevens, K.N., Lindblom, B. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 1003–1010.

Levy, E.S., Strange, W. 2008. Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics* 36, 141–157.

McAllister, R., Flege, J., Piske, T. 2002. The influence of the L1 on the acquisition of Swedish vowel quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, 30, 229–258.

Peltola, M.S., Tuomainen, O., Koskinen, M., Aaltonen, O. 2007. The effect of language education on the preattentive perception of native and non-native vowel contrasts. *Journal of Psycholinguistic Research*.

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., Kuntola, M., Tamminen, H., Hämäläinen, H., Aaltonen, O. 2005. Early exposure to non-native language alters preattentive vowels discrimination. *Neuroscience Letters*, 388: 121–125.

Strange, W. 2007 Cross-language phonetic similarity of vowels: Theorertical and methodological issues. In O.S. Bohn and M.J. Munro (eds.) Language experience in second language speech learning. John Benjamins, Amsterdam, 35–56.

Trofimovich, P., Baker, W., Mack, M. 2001. Context- and experience-based effects on the learning of vowels in a second language. *Studies in the Linguistic Sciences, 31*, 167–186.

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J.E. 2005. A developmental study of English vowel production and perception by native English adults and children. *Journal of Phonetics*, 33, 263–290.

Weber, A., Cutler, A. 2004. Lexical competition in non-native spoken- word recognition. *Journal of Memory and Language*, 50, 1–25.

Werker, J.F. 1994. Cross-language speech perception: Development does not involve loss. In J.C. Goodman & H.C. Nusbaum (eds.), *The development of speech perception*. MIT Press, Cambridge, MA, 93–120.

Werker, J.F., Tees., R.C. 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behaviour and Development*, 7, 49–63.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csépe, V., Näätänen, R. 1999. Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, 26, 638–642.