



HAL
open science

Neurophysiology of non-native sound discrimination: Evidence from German vowels and consonants in successive French–German bilinguals using an MMN oddball paradigm

Jane Wottawa, Martine Adda-Decker, Frédéric Isel

► **To cite this version:**

Jane Wottawa, Martine Adda-Decker, Frédéric Isel. Neurophysiology of non-native sound discrimination: Evidence from German vowels and consonants in successive French–German bilinguals using an MMN oddball paradigm. *Bilingualism: Language and Cognition*, 2021, pp.1-11. 10.1017/S1366728921000468 . hal-03321725

HAL Id: hal-03321725

<https://hal.science/hal-03321725>

Submitted on 5 Jan 2022

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

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

Running head:

Neurophysiology of non-native German sound discrimination

Title:

Neurophysiology of non-native sound discrimination: Evidence from German vowels and consonants in successive French-German bilinguals using an MMN oddball paradigm*

^{1,2}Jane Wottawa, ²Martine Adda-Decker, ³Frédéric Isel

¹*Le Mans University, LIUM*

²*Sorbonne Nouvelle – Paris 3 University – CNRS, UMR 7018, LPP*

³*Paris Nanterre – Paris Lumières University, CNRS, UMR 7114, MoDyCo*

*We'd like to thank the Max-Planck-Institute for Human Cognitive and Brain Sciences in Leipzig (Germany) for allowing us to collect data there and to exchange with its members. We'd further like to thank Cornelia Henschel and Mahsa Bahrami who helped with the recruitment of participants and data collection. This work was funded by the French *Investissements d'Avenir - Labex EFL* program (ANR-10-LABX-0083).

Address for correspondence:

Jane Wottawa,

Le Mans Université

UFR de Lettres

Avenue Olivier Messiaen

72085 Le Mans (France)

jane.wottawa@univ-lemans.fr

Keywords:

L2 German sound discrimination, EEG, MMN, oddball experiment, vowels, fricatives

Abstract (<150 words)

The present electroencephalographical multi-speaker MMN oddball experiment was designed to study the phonological processing of German native and non-native speech sounds.

Precisely, we focused on the perception of German /ɪ-i:/, /ɛ-ɛ:/, /a-a:/ and the fricatives [ʃ] and [ç] in German natives (GG) and French learners of German (FG). As expected, our results showed that GG were able to discriminate all the critical vowel contrasts. In contrast, FG despite their high L2 proficiency level, were only marginally sensitive to vowel length variations. Finally, neither GG nor FG discriminated the opposition between [ʃ] and [ç], as revealed by the absence of MMN response. This latter finding was interpreted in terms of low perceptual salience. Taken together, the present findings lend partially support to the Perceptual Assimilation Model for late bilinguals (PAM-L2) for speech perception of non-native phonological contrasts.

1 Introduction

Late L2 learners face challenges with respect to speech sound perception, especially for phones that do not exist in the native phonological system. Their well-established phonological system might help or hinder the perception of sounds in their second language (Troubetzkoy, 1949; Flege, 1995; Best, 1996; Schertz, Cho, Lotto & Warner 2015). Models of perception and production in second languages enable the prediction of possible difficulties that can be encountered by second language learners. These models formalize possible interferences between L1 and L2 speech perception and production. They account for typological differences of the learners' L1 and L2 as predictions are made by comparing the phonological systems of the two languages. Unfortunately, most models do not consider learner proficiency on a continuum, consequently the same predictions are usually made for learners who are either at a beginner's or at an advanced level.

Although two major models predicting production and perception biases in late learners of second languages are widely used, namely Flege's Speech Learning Model (SLM; Flege, 1995) and Best's Perceptual Assimilation Model for late bilinguals (PAM-L2; Best & Tyler, 2007), we will focus only on the latter model in the following. The reason for this choice is that only Best's model predicts speech perception of non-native phonological contrasts specifically. The model was inspired by motor-theory which implies that successful perception is linked to successful production of the sound. The predictions of PAM-L2 are based on comparisons between the phonological contrasts existing in the listeners' L1 and L2. Five distinct predictions are made by the model (i.e., *two-category assimilation*, *single-category assimilation*, *category goodness difference*, *uncategorized-categorized assimilation*, and *uncategorized assimilation*).

In order to apply the PAM-L2 model to the German-French language pair and to decide which predictions summarize their phonological differences, a comparison of the respective phonemic inventories was undertaken.

Comparing the phonemic systems of German and French indicates, on one hand, that the vowel systems of German and French differ with respect to their number of vowels. German counts 16 monophthongs not considering the 3 diphthongs, all oral vowels (Kohler, 1999; Mangold & Dudenredaktion, 2005). The French language counts 11 monophthong oral vowels and 3 nasal vowels (Fougeron & Smith, 1999).

< Insert Figure 1 about here >

If we compare the F1-F2 vowel space of the German and the French vowel systems (see Figure 1), we observe that the vowel triangle in German shows a greater density. This difference in vowel space density is linked to the German vowel length contrast which is absent in French. The vowel length contrast in German oral vowels is widely stated throughout the literature (Antoniadis & Strube, 1984; Becker, 1998; Wiese, 2000; Hall, 2011; Ramers, 2012). With respect to phonology, two vocalic features are engaged in this contrast: vowel quality and vowel duration. Pure phonological analyses demonstrate that vowels [+tense] always associate the feature [+long], whereas vowels [-tense] can associate both [-long] and [+long] (e.g., /ɛ:/, a:/) (Hall, 2011). These observations led Ternes (2012) to the conclusion that the primary feature opposing German vowels is vowel length. In the following, we will distinguish between short and long vowels. The differences in oral vowel inventory sizes, vowel space density and the vowel length contrast raised the question of how French learners of German discriminate the German vowel length contrast in speech perception.

On the other hand, the consonantal systems of German and French show a high number of similarities. Fougeron and Smith (1999) established a list of 21 French phonemic consonants which is quantitatively similar to the 20 German phonemic consonants listed by Kohler (1999). However, some of the German consonants are not present in the French consonantal system. French lacks the glottal consonants [ʔ] and /h/, the palatal fricative [ç] (*ich-Laut*) and the uvular fricative [x]¹ (*ach-Laut*). Conversely, some of the French

consonants such as the palatal nasal consonant /ɲ/ are not found in the German consonantal system.

The German fricatives [ç] and [x] although acoustically and perceptually clearly distinct are considered to be allophones. In the German lexicon, they appear in complementary word positions according to the left vowel context. Anterior vowels (as well as some consonants) are followed by [ç], whereas posterior vowels and the vowel /a/ are followed by [x]. Both fricatives appear generally in word final (*Buch* [bu:x], book) or in morpheme final positions (*riech-en* [ʁi:.çən], to smell). In loanwords however, [ç] can also appear word initially (*Chemie* [çɛmi:], chemistry), *China* [çina], China etc.)

In production tasks, French learners of German often replace the consonant [ç] (like in *mich* [miç] - me) by its closest neighbor, the post alveolar fricative /ʃ/ (like in *misch* [miʃ] – to mix) (Wottawa, Adda-Decker, Isel, 2016). The post alveolar fricative /ʃ/ appears syllable initially as well as syllable finally in German: *schnell* [ˈʃnɛl] (fast) vs. *Fisch* [ˈfiʃ] (fish), and in the adjectival suffix *-isch*. At the end of monosyllabic words, the voiceless palatal fricative [ç] often appears in a cluster with the plosive [t]: *Licht* [ˈliçt] (light), *echt* [ˈɛçt] (real). Regarding the derivational morphology of German, the voiceless palatal fricative [ç] appears in the suffixes *-chen* and *-(l)ich* as well.

From an acoustic point of view, the German [ʃ] and [ç] are highly similar. Their respective points of articulation are very close to each other even if the articulatory movements are quite different. The respective centres of gravity of the two fricatives are situated in similar frequency bands. These similarities are particularly critical with respect to the question of non-native contrasts perception.

As we were interested in the differences between French and German, we applied the PAM-L2 model to these two languages. The articulatory dimension of the model was not investigated further as we did not record participants' articulatory data in this study. We reduced the model on the phonemic categories of the participants L1 and L2. In comparing

the phonological systems of German and French through the looking glass of PAM-L2, it appears that French learners of German are mainly confronted with single category assimilation (i.e., two distinctive non-native sounds are assimilated as variants of the same category) and category goodness difference (i.e., two distinctive non-native sounds are assimilated as more or less valid prototypes of the same category).

On one hand, the perception of the German vowel length contrast in French learners of German seems to belong to the *category goodness difference*. With respect to vowel length and vowel quality, in isolated spoken words, French vowels are more similar to long vowels than to short vowels of German (Strange, Weber, Levy, Shafiro, Hisagi & Nishi, 2007). German long vowels should therefore be perceived as good exemplars of the vowel category in French native listeners, whereas short vowels might be perceived as poor(er) examples of this category. Nevertheless, the acoustic differences might not be salient enough to French learners of German to perceptually separate short from long vowels. In our study, the German vowel pairs differ all in duration in a similar way, however spectral properties vary to different degrees: the /ɪ-i:/ pair shows considerable differences in vowel quality, whereas /ɛ-ε:/ and /a-a:/ have quite similar vowel qualities. The evaluation of category-goodness might be influenced by the identity of the investigated vowel pair. In that case, German [ɪ] should be perceived as a poor candidate for the French learners' /i/ category, whereas German [i:] should be a good candidate for this category. German [ɪ] presents different articulatory movements of the jaw and tongue but especially the lips than does German [i:]. The German vowels [a] and [a:] should be equally good candidates for the French learners' /a/ category as the articulatory movements are highly similar and the vowels differ mainly in duration, which should increase the perception difficulty in French native listeners.

On the other hand, the perception of the German consonants [ʃ] and [ç] seems to belong to the *single category assimilation* scenario. Especially as the fricative /ʃ/ is part of the French phonemic inventory. Late French-German bilinguals already have a representation for

/f/ with an automated articulatory gesture. The acoustic properties of German [f] and [ç] are very similar (Wottawa et al., 2016; Jannedy & Weirich, 2016) whereas the articulatory movements particularly of the tongue and the lips are quite different. Based on the acoustic information, it is unlikely that L2 learners of German whose L1 presents a relatively low number of fricative types consider the unknown but similar [ç] as an uncategorized sound. From an articulatory point of view, it is possible that even advanced French learners of German did not yet experience the canonical articulatory movement of [ç] because deducing a different articulatory movement from an acoustically similar sound is a very hard task for late learners with an already well-established phonological system. Thus, we assume that learners in our group assimilate [f] and [ç] into one category. The PAM-L2 predicts rather low or no discrimination for sound contrasts that are assimilated by one single L1 phonemic category.

There are different methods that allow us to investigate non-native speech perception such as behavioral perception tests and neural imaging methods. We chose to investigate the perception of the German vowel length contrast and the opposition of [f] and [ç] using electroencephalography (EEG). EEG allows us to record brain responses millisecond by millisecond, and, in some settings, without the participant needing to perform an experimental task. Combining EEG recordings with a Mismatch negativity oddball paradigm, which is assumed to sign phonological discrimination processes (Dehaene-Lambertz, 1997) without involvement of attentional resources, ensured the ability to tap the automaticity of the processes by minimizing interferences, among others, of motor or decision processes (Winkler, Lehtokoski, Alku, Vainio, Czigler, Csépe, Aaltonen, Raimo, Alho, Lang, Iivonen & Näätänen, 1999). Hence, this technique enables us to better understand early processing of sound discrimination in native and non-native listeners.

The classical oddball paradigm consists of a number of stimuli sequences combining two types of stimuli: similar (standard) and “odd” (deviant) ones. Deviant stimuli are rare (i.e., 10 %) compared to the frequent standard stimuli (i.e. 90 %) (Remington, 1969; Squires,

Wickens, Squires & Donchin, 1976). The Mismatch negativity (MMN) (Näätänen, Gaillard & Mäntysalo, 1978) is observed on deviant stimuli varying in some acoustic property from standard stimuli. The MMN is the difference of two ERP wave forms: the averaged wave forms associated with the deviant sounds (deviant condition) and the averaged wave forms elicited by the processing of the same stimulus but in a standard condition (i.e., preceded by the same sound). The subtraction of the standard condition, whose trace is kept in sensory memory (Cowan, Winkler, Teder & Näätänen, 1993; Näätänen & Winkler, 1999; Picton, Alain, Otten, Ritter & Achim, 2000; Winkler & Cowan, 2005), from the deviant one is expected to result in an MMN peaking around 100-200 ms after stimulus onset, if any. The MMN reflects the early process of auditory novelty detection (Schröger, 1998; Näätänen, 2018) and is elicited by the auditory cortex (Alho, 1995; Rinne, Alho, Ilmoniemi, Virtanen & Näätänen, 2000). Recently, the MMN was also found to reflect variations of phonetic information related to vowels conveying different emotional information (Carminati, Fiori-Duharcourt & Isel, 2018). Its maximum amplitude is observed in the fronto-central regions of the scalp.

The MMN is used as a marker in L2 studies in order to investigate the discrimination capacities of L2 contrasts in second language learners. The emergence of MMN as a function of linguistic experience constitutes a good criterion of neuroplasticity. Most of the time, auditory stimuli are employed in order to investigate phonetic or phonological categories in L2 listeners i.e., Catalan vowels (Díaz, Baus, Escera, Costa & Sebastián-Gallés, 2008); English vowels (Peltola, Kujala, Tuomainen, Aaltonen & Näätänen, 2003; Shafer, Yan & Datta, 2011; Grimaldi, Gili Fivela, Invitto, Resta, Alku & Brattico, 2014; García & Froud, 2018; Krzonowski, Pellegrino, & Ferragne, 2018); Finnish vowels (Nenonen, Shestakova, Huotilainen & Näätänen, 2005; Savo & Peltola, 2019); German vowels (Rinker, Alku, Brosch & Kiefer, 2010); and English consonants (Iverson, Ekanayake, Hamann, Sennema & Evans, 2008; Mah, Goad & Steinhauer, 2016) but also non-native prosody (Friedrich, Herold, &

Friederici, 2009). Furthermore, it is known that individual differences in L2 listeners such as language proficiency, musicality or the amount of language practice might influence the MMN amplitude (Peltola et al. 2003, Tervaniemi, Ilvonen, Karma, Alho & Näätänen, 1997, Grimaldi et al., 2014).

Under certain conditions, the negative MMN is followed by a positive deflection: the P3a, a variant of the P300 appearing in passive conditions resulting in a biphasic ERP pattern (Squires, Squires & Hillyard, 1975; Snyder & Hillyard, 1976). The P3a-component is thought to reflect attentional resources, whereas the MMN is an automatic reaction to acoustic changes. The P3a is a positive waveform peaking between 220 – 280 ms after stimulus onset, elicited in passive paradigms. This ERP shows its maximum amplitude over fronto-central electrodes. The P3a, also called novelty P3, is interpreted as an involuntary attention shift as a reaction to changes in the environment or the processing of new information (Yamaguchi & Knight, 1991a,b; Fiori-Duharcourt & Isel, 2012).

The present study aims to investigate the perception of some German vowel and fricative oppositions in French learners of German that we expect to be particularly difficult for French listeners. The German vowel length contrast, which has a lexical value (word opposition), appears in numerous minimal pairs e.g., *bitten* ['bitən] (to ask, to solicit) and *bieten* ['bi:tən] (to offer). Regarding the opposition of [ʃ] and [ç], which exists in German (but not in French), only few minimal pairs can be found e.g., *misch* ['miʃ] (to mix) vs. *mich* ['miç] (me). We specifically selected these two oppositions as their frequency of occurrence in the German language is rather different. Whereas vowel length oppositions are a frequent phenomenon in German, it is less the case for the opposition of [ʃ] and [ç].

Our goal was to investigate to what extent the phonological system of native listeners of French and advanced learners of German is sensitive to the phonological properties of the German language. Sensitivity to phone contrasts of German that do not exist in the phonemic register of French observed in adult second language learners is a strong argument that the

phonological processing system has sufficient deformability to adapt to the new sounds encountered in a second language (Perani & Abutalebi, 2005, Costa & Sebastián-Gallés, 2014). Empirical evidence of such adaptation abilities might be reflected at the neural level by neuronal changes, particularly in terms of neurophysiological responses.

We approached the question of cognitive plasticity at the phonological level of processing by using a neurophysiological marker, i.e., the MMN, which is thought to be a relevant passive sound discrimination signature in psycholinguistic and phonetic literature (Winkler et al., 1999). Thus, in order to investigate the perception of the two contrasts in native and non-native learners of German, an EEG experiment with a passive oddball paradigm was designed. Discrimination performances of advanced German learners were compared to German native listeners. Variations of amplitude, peak latency, and surface topography (scalp distribution of the electrodes) of the MMN constitute critical dimensions (Zevin, Datta, Maurer, Rosania & McCandliss, 2010, Díaz, Mitterer, Broersma, Escera & Sebastian-Galles, 2016) to study auditory processing differences between native speakers and late bilinguals.

Based on the predictions of the PAM-L2 model, different neural responses according to the studied L2 phone contrasts were expected. Vowels which have both, spectral and duration differences e.g., the vowel pair [ɪ-i:] are expected to be better discriminated in French learners of German than vowel pairs that have little spectral variation e.g., [ɛ-ɛ:] and [a-a:]. A better perception of the acoustic differences between members of a given vowel pair should elicit a higher amplitude in both the MMN and P3a. Moreover, processing of deviant stimuli latencies should lead to a latency shortening of the biphasic ERP pattern MMN/P3a for [ɪ-i:], whereas [ɛ-ɛ:] and [a-a:] should not provoke any latency shortening. According to PAM-L2, no or little discrimination is expected for the opposition of [ʃ] and [ç] in non-native listeners which might result in an MMN and P3a with low amplitudes.

In the following, we'll present first the applied Methods, followed by the Results section, finally we'll present the Discussion and Conclusion.

2 Methods

In this section, the applied methods and procedures including the choice of participants and stimuli followed by a description of the procedures are presented.

2.1 Participants

Twenty native speakers of French who started learning German at school (mean age of learning: 12.0 years, SD: 1.1 years) were tested in Paris (France). In addition, twenty native speakers of German were recruited in Leipzig (Germany). The French participants were aged from 19 to 34 years (mean age: 22.8 years, SD: 4.3 years). The German participants were aged from 21 to 28 years (mean age: 24.4 years, SD: 2.3 years). All participants had to fit the following criteria: 1) to have no other first language than respectively German (native speaker group) or French (learners' group); 2) to be aged between 18 and 35 years; 3) to be right handed (Edinburgh Handedness Inventory); and 5) to present normal hearing. By their own account, participants had no history of current or past neurological or psychiatric diseases, they had normal or corrected-to-normal vision. Furthermore, the native speakers of German had no or very limited knowledge of Romance languages i.e., no or only an introductory class during high school while the French participants had to have regular contact with German (mean = 18.6h / week) during the past year through language classes or a teaching activity, for instance. The participants gave written consent after obtaining all necessary information on the experiment as well as data storage and anonymization. The collected data were anonymised by applying the European Data FAIR principle. The study was approved by the local ethics committee of the Paris Nanterre University and was performed in accordance with the Declaration of Helsinki. All EEG participants received 20€ for their time and effort.

2.2 Stimuli

Most oddball paradigms present stimuli that are either synthesized or coming from one speaker alone. This type of stimuli creates a homogeneous acoustic environment where listeners' attention is easily drawn to any changes both acoustically or phonologically (Winkler, 2003). In order to test sensitivity to higher order regularities than mere acoustic changes, acoustically different stimuli can be used. In the present study, this acoustic variability was achieved using stimuli that were produced by multiple speakers. For this purpose, seven female German native speakers were recorded in a sound-proof room at the *Laboratoire de Phonétique et Phonologie – Paris 3*. Only female speakers were chosen in order to avoid reactions to gender change during the oddball EEG experiment (Casado & Brunelière, 2016). The stimuli were isolated words, that were only harmonized for loudness. We postulate that natural speech provides a more realistic condition to investigate human auditory sensitivity to L1 and L2 speech than do synthesized speech or mono-speaker stimuli, especially in the case of a laboratory study.

Vowel duration contrast

The chosen vowel pairs were /ɪ-i:/, /ɛ-ε:/ and /a-a:/ . Only natural words, either monosyllabic or bi-syllabic, were recorded. The target vowel always appeared in stressed syllables. All three vowel pairs differ with respect to their duration: short vowels presented shorter durations than did long vowels. The spectral differences across the three vowel pairs, however, were not similar. The following hierarchy can be established going from the vowel pair with the highest spectral differences to the vowel pair with the lowest spectral differences: /ɪ-i:/ > /ɛ-ε:/ ≥ /a- a:/ (Strange, Bohn, Trent & Nishi, 2004). Spectral differences with respect to vowel discrimination might be of advantage in non-native perception (McAllister, Flege & Piske, 2002).

Figure 2 illustrates an example of a minimal pair containing the target vowels [ɪ-i:]. On top, a spectrogram of the word containing the short vowel is depicted, whereas below the word containing the long vowel is illustrated. The dotted red line in the spectrogram marks the second formant of the target vowel. Figure 2 shows a duration difference between short and long vowels. The spectral differences, in the illustrations are mainly marked by F2 formant values for [ɪ-i:]: e.g., *bitte* (i.e., 1869 Hz) and *biete* (i.e., 2370 Hz).

< Insert Figure 2 about here >

Opposition of [ʃ] and [ç]

Both fricatives [ʃ] and [ç] appear mainly syllable-finally in German *misch* [mɪʃ] (to mix), *mich* [mɪç] (me). There are only few natural minimal pairs in German contrasting [ʃ] and [ç]. Therefore, we added pseudo-words containing this contrast in order to enhance the number of stimuli. It is known that pseudo-words elicit a smaller MMN response than do natural words however pseudo-words, which present valid phonotactics of the target language, are a good compromise to investigate contrasts that exists only in few natural minimal pairs (Pulvermüller & Shtyrov, 2006). The target consonants were recorded in different valid vowel contexts (i.e., [ɪ, ʏ, y:, ε, œ, aɪ, ɔɪ]) and appeared both in sequence internal and in sequences final positions. Two examples of minimal pairs are illustrated in Figure 3. The figure displays typical differences in F2 of the preceding vowels due to the place of articulation of the respective fricatives. Both the preceding vowels as well as the fricative noise (important friction noise for [ʃ], less friction noise for [ç]) were criteria for stimuli selection.

< Insert Figure 3 about here >

2.3 Procedures

The experiment followed the design of a classical passive auditory oddball paradigm. Participants were comfortably seated in front of a computer screen in a sound isolated room

and watched a silent movie. The stimuli were presented over headphones. At the end of the experiment, the participants filled out a questionnaire relative to the movie in order to make sure that their attention was focused on the movie and not on the sounds. The stimuli were organized in 5 sequential blocks. Each vowel contrast (/i-i:/, /ε-ε:/, /a- a:/) was presented in one separate block, furthermore two blocks presented the [ʃ] and [ç] opposition, one for stimuli presenting [ʃ] and [ç] in mid-sequence, and one presenting the fricatives in sequence final positions. The blocks were presented in a randomized order. In order to allow pairwise comparisons, standards could become deviants and vice versa. For instance, the word *biete* was a standard followed by another standard *biete* or the deviant *bitte* as well as the word *bitte* was a standard followed by another standard *bitte* or the deviant *biete*. Moreover, half of the trials² were pure standard trials.

Each trial contained a variable number of items produced by three or more different speakers. In each trial, the stimuli coming from the same speaker were at least separated by two stimuli from different speakers in order to ensure acoustic variability. In the literature, the number of preceding standard stimuli varies from two to eleven (Rosburg, Trautner, Ludowig, Schaller, Kurthen, Elger & Boutros, 2007; Garrido, Friston, Kiebel, Stephan, Baldeweg & Kilner, 2008; Kirmse, Ylinen, Tervaniemi, Vainio, Schröger & Jacobsen, 2008; Biedermann, De Lissa, Mahajan, Polito, Badcock, Connors, Quinto, Larson & McArthur, 2016). Authors of the different studies have different approaches: for instance, Rosburg et al. (2007) presented either two, three or four preceding stimuli, whereas Garrido et al. (2008) presented between one and eleven preceding stimuli. In the present experiment, the last item of a trial was preceded by six up to nine standard stimuli separated by a 500 ms inter stimulus interval. This choice was made because firstly, the stimuli were uttered by different speakers. Habituation with this kind of stimuli should take longer than with synthesized or mono-speaker stimuli that present exactly the same acoustic quality; secondly, after each trial, a different standard word was presented, which led us to reject the first two standards of each

trial in the ERP analyses in order to exclude standards, which carry also an MMN due to item change; and finally, Remington (1969) and Falmagne et al. (1975) found that at least five items in a stimuli chain favor effective reaction times in identifying the deviant stimuli. We considered that neurophysiological responses should also benefit from longer standard chains.

EEG recordings

EEG was recorded from 64 channels mounted in an elastic cap. For data analysis, channels were re-referenced to an average reference. Electrode impedances were kept below 25 k Ω . Data were recorded at a sampling rate of 1000 Hz.

In France, recordings were undertaken with the BrainVision PyCorder, the signal was amplified with the BrainVision actiCHamp amplifier. The used EEG caps were actiCAP caps with 64 electrodes. The ground electrode was attached to the sternum. The device records the online EEG signal against an implicit reference generated by the BrainAmp (Brain Products) amplifier.

In Germany, recordings were undertaken with the BrainVisionRecorder (version 1.20.0601), Brain Product. The signal was amplified with the BrainVision BrainAmpDC amplifier. WaveguardTMoriginal caps comprising 64 electrodes were used at the Max-Planck-Institute in Leipzig. The ground electrode was attached to the sternum, the reference electrode was attached to the tip of the nose.

EEG signal processing and data analyses

The recorded EEG signal was processed with the Matlab toolboxes EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014). In order to prepare the EEG signal for the ERP analyses, it was first down sampled to 500 Hz and filtered with a high-pass filter at 0.1 Hz and a low-pass filter at 45 Hz. Afterwards, a bad channel location was carried out, followed by the interpolation of electrodes in order to keep a high

number of trials. Finally, all channels were re-referenced to the mean of all channels of the cap. This choice was made because the device used in France records against an implicit reference generated by the amplifier. In the next step, epochs were extracted from the pre-processed EEG signal for the three midline electrodes Fz, Cz, Pz where MMN is usually observed (Paavilainen, Karlsson, Reinikainen, & Näätänen, 1989; Alho, 1995). The chosen time window was [-100 ms to 700 ms] from stimulus onset. Baseline correction was carried out on 100 ms before stimuli onset, finally, an automatic artefact rejection with a threshold of 70 μ V was performed.

MMN calculation

In order to image the paradigm settings in the MMN calculations, the average of standards included all standard stimuli presented in the trials of the oddball paradigm, except for the first two items per trial, whereas the average of deviant stimuli included all the deviants which were at most one per trial and always situated at the end of a trial. The averaged waveform of the standards was then subtracted from the averaged waveform of the deviants.

For the vowel duration contrast condition, standards of short vowels were compared to deviants of short vowels and standards of long vowels were compared to deviants of long vowels. For the consonant condition, standards containing [ʃ] were subtracted from deviants containing [ʃ], whereas standards containing [ç] were subtracted from deviants containing [ç].

The time windows of the ERPs were determined based on the literature and then adjusted visually based on the data (begin: latency where the waveform crosses the zero-line, end: latency after the main peak's minimum).

3 Results

In this section, results are presented starting with the vowel duration contrast followed by the opposition of [ʃ] and [ç]. We recall, that the MMN is an automatic reaction to acoustic changes, whereas the P3a-component is thought to reflect attentional resources.

3.1 Vowel duration contrast

MMN

For the MMN amplitude analyses, the time window 90-200 ms was fixed. A four-way ANOVA including the factors Vowel pair (/i-i:/, /ε-ε:/, /a- a:/), Vowel length (short, long), Electrode (Fz, Cz, Pz) as within-subject factors, and the factor Group (German natives, French learners of German) as between-subject factor was run. Only the interaction Group × Electrode was significant ($F_{(2, 74)} = 5.60$, $\eta^2 = .009$, $p < .01$). Planned comparisons indicated that the MMN amplitude was significantly larger at Fz ($M = -0.26 \mu\text{V}$, $SD = 1.26 \mu\text{V}$) than on both Cz and Pz ($M = 0.14 \mu\text{V}$, $SD = 1.32 \mu\text{V}$) for the German group ($F_{(1, 357)} = 7.28$, $p < .01$), while the MMN did not differ between the three electrodes in the bilingual French-German group (Fz: $M = -0.08 \mu\text{V}$, $SD = 1.62 \mu\text{V}$; Cz: $M = -0.19 \mu\text{V}$, $SD = 1.85 \mu\text{V}$; Pz: $M = -0.12 \mu\text{V}$, $SD = 1.59 \mu\text{V}$). Furthermore, a one sample t-test comprising the values from Fz, Cz, and Pz against zero revealed that the MMN in the bilingual French-German group was only marginally significant ($t_{(359)} = 1.9$, $p = .05$).

In German native speakers, a one sample t-test comparing the values of Fz against zero was significant ($t_{(119)} = |2.2|$, $p < .05$). Figure 4 illustrates this interaction.

< Insert Figure 4 about here >

With respect to peak latency, a four-way ANOVA including the factors Vowel pair, Vowel length, Electrode as within-subject factors, and the factor Group as between-subject factor was run. Neither main effects nor interactions were found.

P3a

Analyses on the P3a amplitude values, expected to peak between 220 – 280 ms, were extracted for each participant in the time window located between 190 and 240 ms after stimuli onset. No P3a was observed for the two speaker groups.

Late negativity

Interestingly, we found a negative deflection at Pz in the time window situated between 400 and 460 ms after stimuli onset. The negativity's amplitude values for each participant were extracted from the time window situated between 400 and 460 ms after stimuli onset for the Pz electrode. A four-way ANOVA with the within-subject factors Condition (standard, deviant), Vowel pair, Vowel length, and the between-subject factor Group was run. The ANOVA showed a marginally significant main effect of Condition ($F_{(1,37)} = 3.15$, $\eta^2 = .002$, $p = .08$), which led us to carry out separate analyses for each group. According to the assumption that a late negativity in a time window around 400 ms might indicate involvement of lexical access processes, further analyses were carried out for the Pz electrode where the N400 is usually observed. A three-way ANOVA was run for each group comprising the within-subject factors Condition, Vowel pair, Vowel length. A main effect of Condition was observed only in German native speakers ($F_{(1,19)} = 7.57$, $\eta^2 = .02$, $p < .05$) as illustrated by Figure 5. The effect had an amplitude of $-0.30 \mu\text{V}$ ($\text{SD} = 1.55 \mu\text{V}$).

< Insert Figure 5 about here >

With respect to peak latency at the Pz electrode, a four-way ANOVA with Condition, Vowel pair, and Vowel length as within-subject factors and Group as between-subject factor showed no main effects or interactions.

3.2 Opposition of [ʃ] and [ç]

MMN

The MMN amplitude was extracted in the time window situated between 350 and 550 ms after stimulus onset, counting in the delay between the pseudo-word beginning and the start of the target fricative.

A four-way ANOVA including the factors Consonant (ʃ, ç), Position (word-internal, word-final), Electrode (Fz, Cz, Pz) as within-subject factors, and the factor Group (German natives, French learners of German) as between-subject factor was conducted. Neither significant main effects nor interactions were found.

P3a

Analyses on the P3a amplitude values were extracted for each participant in the time window located between 400 and 700 ms after stimuli onset. No P3a was observed for the two speaker groups.

Late negativity

Phone onsets were late thus no late negativity could be observed in the extracted epochs.

4 Discussion

Our goal was to determine to what extent the phonological system of native speakers of French is sensitive to phonological specificities of the German language. To this aim, we investigated the perception of German 1) vowel duration contrasts and 2) of acoustically close consonant oppositions like [ʃ] and [ç] in French learners of German using an EEG auditory oddball multi-speaker experiment. Our motivation to specifically select these contrasts was

that the PAM-L2 model makes different predictions about their processing in French native listeners. Furthermore, we were interested in using sound pairs with different occurrence frequencies in the German language where the vowel length oppositions are a frequent phenomenon in the German lexicon and spoken language, while it is less the case for the fricative opposition of [ʃ] and [ç]. Moreover, whereas the variation of vowel duration has a lexical value (minimal pairs) in German, it is only a phonetic variation in standard French without lexical consequences. For this purpose, twenty native speakers of French and twenty native speakers of German were tested.

Results show that the vowel duration contrast and the opposition of [ʃ] and [ç] were processed in a different way by the participants of the two groups. In German native speakers, an MMN was present for all the tested vowel duration contrasts. Importantly, the spectral differences of the vowel pairs i.e., /ɪ-i:/, /ɛ-ε:/ and /a-a:/ had no impact on their MMN amplitude. This result suggests that at the early stage of auditory processing, German vowel length is sufficient. Vowel duration seems to be a more salient feature than is vowel quality. The vowel length contrast is marked by duration and, depending on vowel type, by spectral differences. The perceived information is therefore acoustically rich and multidimensional. Thus, changes in the stimuli are detected at an early auditory stage.

With respect to the opposition of [ʃ] and [ç], unexpectedly no MMN was found in German native speakers. This observation might be explained by the acoustic properties of the tested stimuli. The [ʃ] and [ç] fricatives only display limited acoustic differences. The respective centres of gravity of the two fricatives are situated in similar frequency bands. [ʃ] presents merely a stronger friction sound than does [ç]. The results suggest that the acoustic-phonetic differences between [ʃ] and [ç] are not salient enough to be detected by the perceptual system at an early stage of speech processing. The latter argument is reflected by an absence of an MMN in both speaker groups.

Critically, with respect to the addressed question in the present study, French learners of German failed to show an MMN in response to the processing of the two tested contrast types. However, and importantly, we found an emerging negativity distributed across the three midline electrodes (i.e., Fz, Cz, Pz) in association with the vowel duration contrast at an early stage (i.e., 90-200 ms). This distributed but emerging negativity might indicate processing difficulties of the otherwise rich acoustic differences between short and long vowels in German at this early auditory stage after stimulus onset. As for German natives, the observed MMN amplitude did not vary according to the tested vowel pairs in non-native listeners. This result suggests that French learners of German have difficulties separating short from long vowels even if spectral differences are present depending on the vowel pair. We recall that the French lexicon does not present minimal pairs opposing short and long vowels. Acoustic variability was introduced in our experiment using stimuli produced by multiple speakers. Interestingly in our experiment, the MMN amplitude in the non-native group was not correlated with their language proficiency, their musicality, or their amount of German practice per week.

In both speaker groups, no P3a could be observed for either of the tested contrasts. A possible explanation for this absence of effect might be that the linguistic stimuli in the multi-speaker experiment did not present a salient enough acoustic difference in order to capture the participants' attention. The P3a-component is thought to reflect attentional resources, whereas the MMN is an automatic reaction to acoustic changes. It is possible that the stimuli elicited an MMN, as for the native group, at least for the vowel duration contrast, but did not trigger the involuntary attention shift assumed to be marked by the P3a. The multi-speaker design might have hindered the involuntary attention shift for deviant stimuli as the acoustic properties of the stimuli changed with each item in the stimuli stream. Thus, the acoustic differences of the vowels and consonants might have been partially masked due to speaker change, at least at an early attentional level.

Importantly, at a later stage of language processing, in German native speakers, our results showed a late negativity in a time window between 400 and 460 ms after stimuli onset for all tested vowel length contrasts. The stimuli used for this experimental condition were natural German word stimuli. Hence, we hypothesize that the stimuli changes triggered lexical access like unrelated word pairs in a priming experiment (e.g., *bitten* ['bɪtən] (to ask, to solicit) - *bieten* ['bi:tən] (to offer)), at least in the native speaker group. The phonemic changes leading to a lexical change might have contributed to creating an incongruent word pair situation in native speakers of German but not in second language learners. The late negativity might be an N400-like component. The standard and deviant stimuli were always existing minimal pairs of the German language which seemed to trigger the N400-like neural response.

In the light of these results, we will now discuss the predictions made by the PAM-L2 model. Our results clearly show that the vowel duration contrast and the [ʃ]-[ç] contrast are not processed in the same way by none of the two groups: German natives and French learners of German. However, the L2-speech models made predictions for L2 learners without taking into consideration the acoustic information that is carried by the different contrasts. Our experiment showed that a rich acoustic contrast such as the German vowel duration contrast elicits an MMN in German natives and an emerging negativity in French learners of German. However, the negativity's amplitude was smaller in the learners' group and displayed a different distribution in comparison with the native speakers (negativity more largely distributed over the midline electrodes in the non-native group). Furthermore, our data indicated that contrasts with small acoustic differences such as the opposition of [ʃ] and [ç] failed to elicit an MMN in both speaker groups.

According to the PAM-L2 model, the vowel duration contrast can be considered as a “category goodness difference” and should present processing differences according to the tested vowel pairs. Contrasting vowel pairs that differ not only in duration but also present

spectral differences (i.e., /ɪ-i:/) should be discriminated more effortlessly than vowel pairs that differ only in duration (i.e., /a-a:/). Our results did not support this hypothesis. French native speakers seem to process more easily the non-native phonological features [-long] and [+long] than the non-native spectral changes in the vowel pairs in order to discriminate the German vowel contrast. Furthermore, the model predicted less successful discrimination of the vowel duration contrast in L2 speakers than in L1 speakers. This prediction was confirmed by our results.

With respect to the opposition of [ʃ] and [ç], the model predicted no or very little discrimination in French learners of German because of the phones' assimilation to one single native phonemic category. This prediction holds, but also meets some limits. Our data showed that both groups lacked an MMN for the fricatives. In the light of the MMN, which indicates early auditory processing mechanisms, we hypothesize that the acoustic differences of [ʃ] and [ç] in a multi-speaker design are not salient enough to trigger an automatic early auditory response in human listeners. Thus, the poor discrimination is not directly linked to the participants' first language but rather to the acoustic properties of the tested contrasts.

Our results suggest that a comparison of L2 system to the phonetic or phonemic inventory of the learners' L1 does not always allow fine-grained predictions of perception difficulties in L2 speakers. Especially with respect to phones that present few acoustic differences and whose contrasting perception is difficult even for native speakers who are exposed to them on a daily basis.

5 Conclusions

The present study showed that the German vowel duration contrast and the consonant opposition involving the fricatives [ʃ] and [ç] are processed in different ways by German native and non-native participants. All tested vowel duration contrasts elicited an MMN in German native speakers and an emerging MMN with a broader midline distribution in French

learners of Germans, whereas the fricative contrast did not elicit an MMN in either group. These results suggest that the non-native speakers in our experiment, although exposed to the German language on a regular basis, had processing difficulties with respect to the German vowel duration contrasts. The absence of the N400-like effect in the learner group indicates also that the vowel duration contrast is not (yet) treated as a phonological difference by the tested non-native participants. Regarding the fricative contrast, our results suggest that the acoustic differences were not salient enough to elicit an automatic auditory response in either listener group, at least in a multi-speaker setting. It seems that the perception of non-native contrasts is not only conditioned by the similarity of the non-native phones to the phones included in the phonetic inventories of the learners' L1 but also by the acoustic properties of the stimuli themselves. Our data indicate that the acoustic differences of [ʃ] and [ç] are not salient enough to elicit an MMN, at least in a multi-speaker experiment.

References

- Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and hearing, 16*(1), pp. 38-51.
- Antoniadis, Z., & Strube, H. W. (1984). Untersuchungen zur spezifischen Dauer deutscher Vokale. *Phonetica, 41*(2), pp. 72-87.
- Becker, T. (1998). *Das Vokalsystem der deutschen Standardsprache*, volume 32. Peter Lang, Europäischer Verlag der Wissenschaften.
- Best, C. T. (1996). Non-native speech perception as a window on the native phonological system and its development. *Acoustical Society of America Journal, 100*, 2727.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. *Language experience in second language speech learning: In honor of James Emil Flege, 1334*, pp. 1-47.
- Biedermann, B., De Lissa, P., Mahajan, Y., Polito, V., Badcock, N., Connors, M. H., Quinto, L., Larson, L., & McArthur, G. (2016). Meditation and auditory attention: An ERP study of meditators and non-meditators. *International Journal of Psychophysiology, 109*, pp. 63-70.
- Carminati, M., Fiori-Duharcourt, N., & Isel, F. (2018). Neurophysiological differentiation between preattentive and attentive processing of emotional expressions on French vowels. *Biological Psychology, 132*, pp. 55-63.
- Casado, A., & Brunellière, A. (2016). The influence of sex information into spoken words: a mismatch negativity (MMN) study. *Brain research, 1650*, pp. 73-83.
- Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain?. *Nature Reviews Neuroscience, 15*(5), pp. 336-345.

- Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). *Journal of experimental psychology: Learning, Memory, and Cognition*, 19(4), pp. 909-921.
- Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport*, 8(4), pp. 919-924.
- Delattre, P. (1966). *Studies in French and comparative phonetics: selected papers in French and English*, volume 18. Walter de Gruyter GmbH & Co. KG.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1), pp. 9-21.
- Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences*, 105(42), pp. 16083-16088.
- Díaz, B., Mitterer, H., Broersma, M., Escera, C., & Sebastian-Galles, N. (2016). Variability in L2 phonemic learning originates from speech-specific capabilities: An MMN study on late bilinguals. *Bilingualism: Language and Cognition*, 19(5), pp. 955-970.
- Falmagne, J. C., Cohen, S. P., & Dwivedi, A. (1975). Two-choice reactions as an ordered memory scanning process. *Attention and performance V*, pp. 296-344.
- Fiori-Duharcourt, N., & Isel, F. (2012). *Les neurosciences cognitives*. Armand Colin.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, pp. 233-277.
- Fougeron, C. & Smith, S. (1999). Illustrations of the IPA: French. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*, pp. 78-81.

- Friedrich, M., Herold, B., & Friederici, A. D. (2009). ERP correlates of processing native and non-native language word stress in infants with different language outcomes. *Cortex*, 45(5), pp. 662-676.
- García, P. B., & Froud, K. (2018). Perception of American English vowels by sequential Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 21(1), pp. 80-103.
- Garrido, M. I., Friston, K. J., Kiebel, S. J., Stephan, K. E., Baldeweg, T., & Kilner, J. M. (2008). The functional anatomy of the MMN: a DCM study of the roving paradigm. *Neuroimage*, 42(2), pp. 936-944.
- Grimaldi, M., Sisinni, B., Gili Fivela, B., Invitto, S., Resta, D., Alku, P., & Brattico, E. (2014). Assimilation of L2 vowels to L1 phonemes governs L2 learning in adulthood: a behavioral and ERP study. *Frontiers in human neuroscience*, 8, 279, pp. 1-14.
- Hall, T. A. (2011). *Phonologie: Eine Einführung*. Walter de Gruyter.
- Iverson, P., Ekanayake, D., Hamann, S., Sennema, A., & Evans, B. G. (2008). Category and perceptual interference in second-language phoneme learning: An examination of English/w/-/v/learning by Sinhala, German, and Dutch speakers. *Journal of experimental psychology: human perception and performance*, 34(5), pp. 1305-1316.
- Jannedy, S., & Weirich, M. (2016). The acoustics of fricative contrasts in two German dialects. *Proceedings of Phonetics and Phonology in German Speaking Areas*, pp. 70-73.
- Kirmse, U., Ylinen, S., Tervaniemi, M., Vainio, M., Schröger, E., & Jacobsen, T. (2008). Modulation of the mismatch negativity (MMN) to vowel duration changes in native speakers of Finnish and German as a result of language experience. *International Journal of Psychophysiology*, 67(2), pp. 131-143.
- Kohler, K. (1999). Illustrations of the IPA: German. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*, pages 86–89.

- Krzonowski, J., Pellegrino, F. & Ferragne, E., 2018, Etude acoustique de la production de voyelles de l'anglais par des apprenants francophones, *Proc. Journées d'études de la parole*, Aix-en-Provence (France), 4-8 June, pp. 523-531
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in human neuroscience*, 8, 213, pp. 1-14.
- Mah, J., Goad, H., & Steinhauer, K. (2016). Using event-related brain potentials to assess perceptibility: The case of French speakers and English [h]. *Frontiers in Psychology*, 7, 1469, pp. 1-14.
- Mangold, M. & Dudenredaktion (2005). *Duden: das Aussprachewörterbuch*. Der Duden in 12 Bänden: Das Standardwerk zur deutschen Sprache. Dudenverlag.
- McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of phonetics*, 30(2), pp. 229-258.
- Näätänen, R. (2018). *Attention and brain function*. Routledge.
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological bulletin*, 125(6), pp. 826-859.
- Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta psychologica*, 42(4), pp. 313-329.
- Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2005). Speech-sound duration processing in a second language is specific to phonetic categories. *Brain and language*, 92(1), pp. 26-32.
- Paavilainen, P., Karlsson, M. L., Reinikainen, K., & Näätänen, R. (1989). Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalography and clinical neurophysiology*, 73(2), pp. 129-141.

- 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(1), pp. 25-28.
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current opinion in neurobiology*, 15(2), pp. 202-206.
- Picton, T. W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch negativity: different water in the same river. *Audiology and Neurotology*, 5(3-4), pp. 111-139.
- Pulvermüller, F. & Shtyrov, Y. (2006). Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. *Progress in neurobiology*, 79(1), pp. 49-71.
- Ramers, K. H. (2012). *Vokalquantität und-qualität im Deutschen* (Vol. 213). Walter de Gruyter.
- Remington, R. J. (1969). Analysis of sequential effects on choice reaction times. *Journal of experimental psychology*, 82(2), pp. 250-257.
- Rinker, T., Alku, P., Brosch, S., & Kiefer, M. (2010). Discrimination of native and non-native vowel contrasts in bilingual Turkish–German and monolingual German children: Insight from the Mismatch Negativity ERP component. *Brain and language*, 113(2), pp. 90-95.
- Rinne, T., Alho, K., Ilmoniemi, R. J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *Neuroimage*, 12(1), pp. 14-19.
- Rosburg, T., Trautner, P., Ludowig, E., Schaller, C., Kurthen, M., Elger, C. E., & Boutros, N. N. (2007). Hippocampal event-related potentials to tone duration deviance in a passive oddball paradigm in humans. *Neuroimage*, 37(1), pp. 274-281.

- Savo, S., & Peltola, M. S. (2019). Arabic-speakers Learning Finnish Vowels: Short-term Phonetic Training Supports Second Language Vowel Production. *Journal of Language Teaching and Research*, 10(1), pp. 45-50.
- Schertz, J., Cho, T., Lotto, A. & Warner, N. (2015). Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of phonetics*, 52, pp. 183-204.
- Schröger, E. (1998). Measurement and interpretation of the mismatch negativity. *Behavior Research Methods, Instruments, & Computers*, 30(1), pp. 131-145.
- Shafer, V. L., Yan, H. Y., & Datta, H. (2011). The development of English vowel perception in monolingual and bilingual infants: neurophysiological correlates. *Journal of Phonetics*, 39(4), pp. 527-545.
- Snyder, E., & Hillyard, S. A. (1976). Long-latency evoked potentials to irrelevant, deviant stimuli. *Behavioral Biology*, 16(3), pp. 319-331.
- Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, 193(4258), pp. 1142-1146.
- Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and clinical neurophysiology*, 38(4), pp. 387-401.
- Strange, W., Bohn, O. S., Trent, S. A., & Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *The Journal of the Acoustical Society of America*, 115(4), pp. 1791-1807.
- Strange, W., Weber, A., Levy, E. S., Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *The Journal of the Acoustical Society of America*, 122(2), pp. 1111-1129.

- Ternes, E. (2012). *Einführung in die Phonologie*. Die Sprachwissenschaft. WBG, (Wiss. Buchges.), [Abt. Verlag].
- Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K., & Näätänen, R. (1997). The musical brain: brain waves reveal the neurophysiological basis of musicality in human subjects. *Neuroscience letters*, 226(1), pp. 1-4.
- Troubetzkoy, N. S. (1949). *Principes de phonologie*. Librairie Klincksieck.
- Wiese, R. (2000). *The phonology of German*. Oxford University Press on Demand.
- Winkler, I. (2003). Change detection in complex auditory environment: beyond the oddball paradigm. In *Detection of Change*, pp. 61-81. Springer, Boston, MA.
- Winkler, I., & Cowan, N. (2005). From sensory to long-term memory: evidence from auditory memory reactivation studies. *Experimental psychology*, 52(1), pp. 3-20.
- Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csépe, V., Aaltonen, O., Raimo, I., Alho, K., Lang, H., Iivonen, A., & Näätänen, R. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, 7(3), pp. 357-369.
- Wottawa, J., Adda-Decker, M., & Isel, F. (2016). Putting German [ʃ] and [ç] in two different boxes: native German vs L2 German of French learners. In *Interspeech 2016*, pp. 195-199.
- Yamaguchi, S. & Knight, R. T. (1991a). Age effects on the P300 to novel somatosensory stimuli. *Electroencephalography and clinical neurophysiology*, 78(4), pp. 297-301.
- Yamaguchi, S. & Knight, R. T. (1991b). P300 generation by novel somatosensory stimuli. *Electroencephalography and clinical neurophysiology*, 78(1), pp. 50-55.
- Zevin, J. D., Datta, H., Maurer, U., Rosania, K. A., & McCandliss, B. D. (2010). Native language experience influences the topography of the mismatch negativity to speech. *Frontiers in human neuroscience*, 4, 212.

Figure captions

Figure 1: German and French oral vowel systems of monophthongs. Left: German, (Kohler, 1999); right: French, (Delattre, 1966).

Figure 2: Example of the German vowel length contrast. Stimuli for the vowel pair [ɪ-i:]: top: *bitte* ['bitə] (ask, solicit) and bottom: *biete* ['bi:tə] (offer). F2 is indicated in italic.

Figure 3: Example of the German [ʃ]-[ç] contrast. Stimuli for pseudo words with [ʃ]-[ç] in sequence final position. Top: [gə'pɪʃ], bottom: [gə'pɪç]. F2 is indicated in italic.

Figure 4: MMN of the vowel duration contrast for the three midline electrodes Fz, Cz and Pz in the time window 90–200 ms. A: German natives showed a significant MMN at Fz. B: French learners of German showed an emerging MMN which was distributed over the three midline electrodes.

Figure 5: N400-like effect observed in the time window situated between 400 and 460 ms after stimulus onset for the German vowel duration contrast. A: N400-like effect in German native speakers. B: Absence of the N400-like effect in French learners of German.

¹ The consonants [ʔ], [ç], and [x] are not listed as phonemic consonants by Kohler (1999).

² Each trial corresponds to a series of stimuli that either ends with a standard or a deviant stimulus.