



HAL
open science

Neural evidence for perceiving a vowel merger after a social interaction within a native language

Mélodie Bellegarda, Gary Boddaert, Sophie Dufour, Dominique Knutsen, Angele Brunelliere

► **To cite this version:**

Mélodie Bellegarda, Gary Boddaert, Sophie Dufour, Dominique Knutsen, Angele Brunelliere. Neural evidence for perceiving a vowel merger after a social interaction within a native language. *Brain and Language*, 2025, *Brain and Language*, 261, pp.105529. <10.1016/j.bandl.2024.105529>. <hal-05225371>

HAL Id: hal-05225371

<https://lilloa.hal.science/hal-05225371v1>

Submitted on 27 Aug 2025

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.



HAL Authorization

Title: Neural evidence for perceiving a vowel merger after a social interaction within a native language.

Authors: Mélodie Bellegarda¹, Gary Boddaert², Sophie Dufour³, Dominique Knutsen², & Angèle Brunellière^{2,4}

Affiliations :

¹University of Granada, CIMCYC – Centro de Investigación Mente Cerebro y Comportamiento, 18071 Granada, Spain ; ²Univ. Lille, CNRS, UMR 9193 - SCALab - Sciences Cognitives et Sciences Affectives, F-59000 Lille, France ; ³Aix-Marseille Université, Centre National de la Recherche Scientifique, Laboratoire Parole et Langage, Unité Mixte de Recherche 7309, F-13100 Aix-en-Provence, France; ⁴Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 - CRIStAL, F-59000, Lille, France

Please address correspondence to:

Angèle Brunellière

SCALab, CNRS UMR 9193, Université de Lille, Domaine universitaire du Pont de Bois, BP 60149, 59653 Villeneuve d'Ascq, France

Tel: (+33)3 20 41 72 04

angele.brunelliere@univ-lille.fr

Abstract:

Although previous research has shown that speakers adapt on the words they use, it remains unclear whether speakers adapt their phonological representations, leading them to perceive new phonemic contrasts following a social interaction. This event-related potential (ERP) study investigates whether the neuronal responses to the perception of the /e/-/ɛ/ vowel merger in Northern French speakers show evidence for discriminating /e/ and /ɛ/ phonemes after interacting with a speaker who produced this contrast. Northern French participants engaged in an interactive map task and we measured their ERP responses elicited after the presentation of a last syllable which was either phonemically identical to or different from preceding syllables. There was no evidence for discrimination between /e/ and /ɛ/ phonemes before the social interaction, while mismatch negativity (MMN) and late responses revealed /e/-/ɛ/ discrimination after the social interaction. The findings suggest rapid neuronal adaptations of phonemic representations thanks to the social interaction.

Count: 149 words

Keywords: Speech perception, Phonological variation, Social interaction, Event-related potentials, Mismatch negativity.

1. Introduction

During social interaction in everyday life, speakers are exposed to the large variation that exists in speech production within their native language, such as regional variation from speakers of different geographical regions or variation across speakers' age and gender. It is a considerable challenge to understand how language exposure with such variation shapes the way that the brain treats speech sounds. Behavioral studies have focused on adjustments in production with conversational speech (Pardo, 2006; Pardo et al., 2010, 2013) and on the phenomenon of assimilation in speech production and perception (e.g., Dilley & Pitt, 2007; Gow, 2003). Other studies with behavioral and event-related potentials (ERP) recordings have explored the impact of training with non-interactive settings on the perception of phonemic contrasts (Dufour et al., 2010, 2016; Tang et al., 2020). Despite all these studies in speech production and perception, it remains to determine whether rapid neuronal adaptations at phonological level can emerge just after a social interaction, leading adult speakers to perceive new phonemic contrasts following the interaction. Contrary to non-interactive settings in which linguistic information is not provided in a way that is relevant outside the experimental context, the use of linguistic information is always required for understanding between speakers in social interaction.

Different theories in social interaction (Clark & Brennan, 1991; Giles et al., 1991; Pickering & Garrod, 2004) have posited that individuals adapt their communicative behaviors in such a way as to become more similar to their interlocutor's behavior to increase efficiency in communication between speakers. Among them, Communication Accommodation Theory (CAT, Giles et al., 1991) and interactive alignment model of Pickering and Garrod (2004, 2021) provide interesting proposals on word form adjustment at phonetic/phonological level during a social interaction. According to CAT, individuals adapt to each other's communicative behaviors

under a wider range of phonetic features, including speech rate, intensity and phonological variants. The similarity in communicative acts, called convergence, would guarantee a decrease in social distance between speakers. There is evidence that the convergence can be heightened based on the mutual likability between speakers (Schweitzer & Lewandowski, 2013). In contrast, the interactive alignment model of Pickering and Garrod (2004, 2021) explains the convergent behaviors as the result of the organization of linguistic representations by which an alignment at different levels of representation (semantic, syntactic, and phonological) occurs during the social interaction. In other words, similar representations between speakers at different linguistic levels become activated as the interaction unfolds. However, the two theories make no assumptions as to whether phonological representations are modified following a social interaction and to what extent neural adaptations at the phonological level might be exhibited after a social interaction. In particular, it is not yet known whether the exposure to a new phonemic contrast during an interaction between two native speakers belonging to different regional varieties can shape neuronal discrimination of phonemes following a social interaction.

Phoneme discrimination has been largely studied by measuring the mismatch negativity response (MMN, Näätänen et al., 1978; for a review, Näätänen et al., 2007). MMN is an event-related potential (ERP) elicited by a novel auditory stimulus (usually called deviant stimulus) which interrupted the repeated presentation of another stimulus. This frontocentral negative component peaks at 100–250 ms from the onset of the deviation in the auditory sequence of repeated stimuli. Native phonemic contrasts have been shown to elicit larger MMNs than non-native ones (Dehaene-Lambertz, 1997; Näätänen et al., 1997; Kazanina et al., 2006; Rivera-Gaxiola et al., 2000; Shafer et al., 2004; Winkler et al., 1999) and this effect was observed by the age of 10–12 months (Cheour et al., 1998). Based on these findings, it has been proposed that the

MMN is not only sensitive to acoustic deviance but also to memory traces for native language sounds (Näätänen et al., 2007; Shtyrov & Pulvermüller, 2007).

A large amount of MMN research has focused on neural plasticity for learning new speech sounds or phonetic features before birth (Partanen et al. 2013), in childhood (e.g., Cheour et al., 2002; Shestakova et al., 2003; Peltola et al., 2005) and in adulthood (e.g., Atienza & Cantero, 2001; Kraus et al., 1995; Menning et al., 2000; Näätänen et al., 1993, Tamminen et al., 2015; Tremblay et al., 1997, 1998; Winkler et al., 1999; Zhang et al., 2009). Using a large variety of training methods, including passive exposure, immersive language learning, and intensive training ranging from several days to a two-week period, MMN responses revealed the adaptive brain's ability to discriminate speech sounds due to plastic changes in neural assemblies. For instance, adult Finnish speakers performed listen-and-repeat training on the foreign contrast between /f/ and /v/ English consonants, which does not exist in Finnish (Tamminen et al., 2015). Training effects were found after three days with accurate discrimination on the new phonemic contrast and an increase of the MMN amplitude. Consistent with the crucial role of the superior temporal gyrus in the perception of speech sounds (for reviews, Bhaya-Grossman & Chang, 2022; Yi et al., 2019), the changes in the MMN response related to the discrimination of the /r/-/l/ contrast for adult Japanese speakers after an intensive training were localized in the left hemisphere of the superior temporal region (Zhang et al., 2009).

Most of MMN studies exploring the processing of speech sounds or phonetic features have been conducted with acoustic changes or existing contrasts in foreign languages. In contrast, less attention has been paid to the perception of phonemic contrasts which are subject to variation in production within a native language. Few ERP studies have investigated the

perception of phonemic merger which refers to a contrast for which phonemes are not or are no longer produced in one regional accent but remains distinctively produced in another accent of the same language (Brunellière et al., 2009, 2011; Conrey et al., 2005; Dufour et al., 2013; Tang et al., 2020; Wu et al., 2019). The pioneering ERP study on the perception of phonemic merger (Conrey et al., 2005) found ERP differences in the amplitude of the late positive component (LPC) between /ɪ/-/ɛ/ minimal word pairs (pin/pen) of American English in the unmerged group (i.e., the group producing the /ɪ/-/ɛ/ contrast), whereas there were no such differences in the merged group (i.e., the group who did not distinctively pronounce the phonemic contrast). In accordance with the ERP findings, the merged group had below chance level poor performance for recognizing the /ɪ/-/ɛ/ minimal word pairs, while the unmerged group had good performance. In contrast to the study by Conrey et al. (2005) which presented a prior sentence ending in minimal word pairs, the other ERP studies focusing on phonemic mergers used the classic oddball paradigm with blocks consisting of numerous repetitions of the same auditory stimulus rarely interrupted by another auditory stimulus (Brunellière et al., 2011; Tang et al., 2020; Wu et al., 2019) or an adapted version of the oddball paradigm coupled with a same-different task (Brunellière et al., 2009; Dufour et al., 2013). All these later studies have demonstrated that the phonemic merger affected the MMN response, suggesting that memory traces of speech sounds and spoken words are dependent on the exposure to regional varieties.

To our knowledge, the only ERP study to have examined changes in the neuronal discrimination of a phonemic merger after training showed that merged speakers who did not have the ability to discriminate the Chinese /n/-/l/ contrast could have their memory traces of phonemes rebuilt during the training process (Tang et al., 2020). A MMN response was indeed elicited after a 7-day training where participants received positive feedback on /n/-/l/ contrast

judgements. The training consisted of syllables produced by female and male speakers and participants were asked to select the corresponding letter of the initial phoneme (n or l). It is interesting to note that neuronal discrimination of a new phonemic contrast within the native language was thus found after a training occurring in a non-interactive setting. Therefore, it is not clear whether this could happen after a social interaction, although it is proposed that similar representations between speakers become activated as the interaction unfolds (Pickering & Garrod, 2004, 2021). If speakers' representations become more similar during the interaction, this raises the question of whether changes in memory traces of phonemes happen when a person who does not distinctly produce the two phonemes interacts with a person who does, after the interaction.

In the present study, we seek to examine the neuronal discrimination of the /e/-/ɛ/ merger following a social interaction in Northern French speakers who did not produce and perceive this contrast before the social interaction. The /e/-/ɛ/ merger is a well-known vowel merger in French (Armstrong & Pooley, 2010; Fagyal et al., 2006; Lefebvre, 1991). Although there is a contrastive distinction between /e/ and /ɛ/ in word final open syllables in Standard French, this distinction does not exist in Southern French (Fagyal et al., 2006) or in Northern French (Armstrong & Pooley, 2010; Lefebvre, 1991). For speakers from these regions, the /e/ vowel is produced in that position (e.g., in Northern and Southern French varieties, “*fée*” and “*fait*” are produced with the same phonological form, /fe/, while in Standard French, “*fée*” and “*fait*” are respectively produced /fe/ and /fɛ/). During a social interaction task, participants solved riddles with a Standard French speaking partner in order to find places on a city map and to reach an agreement as to how to name these places (see Figure 1). They had to listen actively to the riddles produced by the Standard French speaking partner in which /e/-/ɛ/ minimal word pairs were inserted and

they then verbally collaborated with their partner to reach an agreement on the name of the places via an audio-conference device. Therefore, participants shared a common goal with their partner.

< Insert Figure 1 >

Before and after the social interaction task (see Figure 1), participants performed a perceptual task which was an adapted version of oddball paradigm coupled with a same-different task (initially developed by Dehaene-Lambertz et al., 2000). Participants heard sequences of three syllables which either were phonemically identical (named control condition) or phonemically differed (named deviant condition) between the final syllable and the two first syllables, and they made a same-different judgment on the last syllable. As in previous studies (Brunellière et al., 2009; Dehaene-Lambertz et al., 2000; Dufour et al., 2013; Eulitz & Lahiri, 2004; Phillips, 2001), each syllable was produced by different speakers in order to force participants to rely on more abstract representations used in the phonological processing. In addition, the syllables were never produced by the Standard French speaking partner with whom the participant interacted to make sure that any neuronal adaptations observed after the social interaction could be interpreted as reflecting the changes of abstract phonemic representations.

We measured the ERP responses elicited by the last syllable in the control and deviant conditions for two phonemic contrasts (/e/-/ɛ/; /i/-/y/) before and after the social interaction. A comparative contrast of the /e/-/ɛ/ merger was selected with the /i/-/y/ contrast which is a stable phonemic contrast across all French-speaking regions. Similar to previous studies on the processing of the /e/-/ɛ/ merger in French (Brunellière et al., 2011; Dufour et al., 2010, 2016), the /e/-/ɛ/ merger should not be distinctly perceived before the social interaction. Then, we hypothesized a neuronal discrimination of the /e/-/ɛ/ contrast after the social interaction over the

MMN response. More precisely, before the social interaction, we predicted a significant interaction between the two contrasts (/e/-/ɛ/ versus /i/-/y/) and the condition factor (control vs. deviant condition) over the MMN, reflecting the lack of ability to discriminate the /e/-/ɛ/ contrast. In contrast, after the social interaction, we predicted a main effect of condition due to the neuronal discrimination of the two contrasts. Based on previous MMN studies on the processing of French phonemic contrasts (Brunellière et al., 2009; Dufour et al., 2013), these effects should be observed on specific topographical sites (left temporal, right temporal, frontocentral, centroparietal, left posterior, right posterior). In line with the ERP data, we predicted a significant interaction between the two contrasts (/e/-/ɛ/ versus /i/-/y/) and the condition factor for the accuracy before the social interaction. Although there should be good performance in both the control and deviant conditions for the /i/-/y/ contrast, a main effect of condition should be found for the /e/-/ɛ/ contrast with good performance in the control condition but below chance level poor performance in the deviant condition due to the lack of ability to differentiate the /e/ and /ɛ/ vowels. Additionally, better performance in the deviant condition for the /i/-/y/ contrast than for the /e/-/ɛ/ contrast should be found. Subsequently, this significant interaction between the two contrasts and the condition factor should disappear after the social interaction with good performance in the deviant condition for the two contrasts.

In addition to the perceptual task, participants performed a reading task before and after the social interaction and we extracted acoustic measures on minimal word pairs differing in either /i/-/y/ contrast or /e/-/ɛ/ contrast for two reasons. The first reason was to demonstrate, prior to the social interaction, a greater Euclidean distance between the values of the F1 and F2 formants of the two vowels for the /i/-/y/ contrast compared with the /e/-/ɛ/ contrast, as well as to identify acoustic overlap for the /e/-/ɛ/ contrast using the Pillai score, a method specifically

employed in the case of vowel mergers (Kelley & Tucker, 2020; Nycz & Hall-Lew, 2013). The second reason was to explore whether participants distinctly produced the /e/-/ɛ/ contrast after the social interaction in the same way as the /i/-/y/ contrast (i.e., similar Euclidean distance between the two contrasts) in accordance with some studies claiming for functional links between speech perception and production (Fadiga et al., 2002; Hickok et al., 2011; Watkins & Paus, 2004).

Moreover, social dimension has been proposed to play a role in convergence (Giles et al., 1991; Pardo et al., 2012; Quezada et al., 2012) and according to Communication Accommodation Theory, (CAT, Giles et al., 1991), individuals adapt to each other's communicative behaviors to achieve a desired social distance between interlocutors. More recently, Gandolfi et al. (2023) have highlighted that social cognition must be considered in the process of building similar representations during the social interaction because individuals would use their interlocutors' point of view to guess what they would say. It thus seems relevant to investigate whether social abilities related to the understanding of the other person's point of view and feeling as defined by the notion of empathy (Cuff et al., 2016) could affect the extent to which phonological representations are modified following a social interaction. More precisely, we addressed an exploratory hypothesis on the impact of empathy by examining whether higher empathy scores assessed by the Interpersonal Reactivity Index questionnaire (French version of IRI, Braun et al., 2015) were associated with more contrastive production and discrimination of the /e/-/ɛ/ contrast after the social interaction.

2. Methods

2.1. Participants

Nineteen (N = 16 females) native French-speaking students (age: $M = 20.47$; $SD = 1.61$) from the University of Lille took part in this study. All were right-handed as assessed by the Edinburgh handedness inventory (Oldfield, 1971); we used the most commonly used 10-item version with two additional items. Participants also reported having no neurological, hearing, visual, or language impairments. Crucially for the aim of this study, they reported being native to the region of Lille in northern France (being born there, growing up there, and currently living there) and having parents from that region¹. They received monetary compensation for participation. The Ethics Committee of the University of Lille approved the experimental procedure used in the study (approval number 2022-588-S104) and each participant provided their written informed consent prior to starting the experiment.

2.2. Materials

2.2.1. Reading Task

¹ While using highly restrictive geographical constraints in the selection of participants allowed us to have a homogenous sample of participants as similar as possible in terms of phonological knowledge, the majority of participants were psychology or speech language therapy students, which prevented us from obtaining a gender-balanced sample. We acknowledge that this can limit the generalizability of our results. As reported in the literature on MMN responses (e.g., Barrett & Fulfs, 1998; Kasai et al., 2002; Toufan et al., 2021), there is currently no consensus as to whether female and male differ from each other; when found, differences when listening to speech input seemed to be limited to auditory processing (Sato, 2020). Instead, the current study sought to explore the adaptation of abstract phonemic representations after a social interaction. However, a series of extensive studies would be needed to understand gender differences in MMN studies by using various stimulus designs (changes in acoustic features, changes in phonemic contrasts with acoustic variability) and tasks (passive and active tasks) and by collecting participants' characteristics such as physiological features, hormonal cycle, and environmental factors. MEG studies with high temporal and good spatial resolution could be expected to be helpful in achieving this goal. However, this is beyond the scope of the current study.

Thirty-six target words were selected from the Lexique database (New et al., 2004), comprised of 18 minimal pairs, half differing in /e/-/ɛ/ phonemic contrast (e.g., “*fée*” /fe/ which means “fairy” in French and “*faits*” /fɛ/ which means “facts” in French), and half differing in the /i/-/y/ control phonemic contrast (e.g., “*ici*” /isi/ which means “here” in French, and “*issue*” /isy/ which means “problem” in French). The difference in each phonemic contrast was always located at the word-final position within an open syllable. All selected words were familiar words with a mean lexical frequency of 100.19 per million words ($SD = 410.80$) and they were either monosyllabic, bisyllabic, or trisyllabic. The members of the minimal pairs within and across the two contrasts were matched in terms of lexical frequency and number of syllables, letters, and phonemes (see, Supplementary Materials). Thirty-six sentences were created, with each sentence containing one of the target words (e.g., the “*fée*” target word in the following sentence “*Sur le dessin, on dirait que la fée prête sa baguette à son amie*” which means “On the drawing, it looks like the fairy is lending her wand to her friend”; and the “*faits*” target word in the following sentence “*Tout le monde parle de ce qu’il s’est passé, mais dans les faits, personne n’a de vraies preuves*” meaning “Everyone’s talking about what happened, but in fact, nobody has any real proof”). The sentences were on average 10.89 words long ($SD = 3.37$), with the target word appearing at varied positions in the sentence. However, the target word was never the first word or the last word of the sentence. It was always followed by a word beginning with a plosive consonant (/d/, /k/, /p/, /t/), which produced a salient event in the speech and reduced its acoustical impact on the final phoneme of the target word, in order to facilitate later acoustic measurements on the final phoneme of the target word. The length of sentences in number of words did not differ between the two phonemic contrasts (see Supplementary Materials).

2.2.2. Perceptual Task

Three native Standard French speakers (two females, one male) produced four syllables several times (/be/, /bɛ/, /bi/, /by/). The best tokens from each speaker were selected as experimental stimuli. By means of visual and auditory inspection using a speech signal editor, we selected the syllables produced by the male speaker that were best matched in fundamental frequency and amplitude envelope because the ERP responses were time-locked to the onset of the last syllable which was produced by the male speaker. All syllables were adjusted to have the same duration of 180 ms (60 ms of prevoicing and 120 ms of plosion and vowel) using Praat software (Boersma & Weenink, 2022). We measured the average Euclidean distance in the F1-F2 plane between the two first syllables and the last syllable at the acoustic mid-point of the vowel: for the /e/-/ɛ/ contrast: $M = 2.03$ Barks; for the /i/-/y/ contrast: $M = 2.34$ Barks. The average distance between the vowels embedded in the two first syllables and the vowel embedded in the last syllable did not differ between the /i/-/y/ and /e/-/ɛ/ contrasts, as revealed by an ANOVA analysis with Condition (control vs. deviant), and Contrast (/e/-/ɛ/ vs. /i/-/y/) as factors (all $ps > 0.2$).

2.2.3. Social Interaction Task

Participants interacted verbally to solve riddles associated with maps with a Standard French speaking partner who was not one of the speakers who had produced the stimuli for the perceptual task. Participants were informed that they interacted with a partner coming from a region other than their own. Five different female speakers² producing the /e/-/ɛ/ contrast played the role of the partner in the study. However, each participant interacted with only one Standard French speaking partner among the five different female speakers.

² The exclusive use of females was made to control for a potential gender bias.

Eleven maps of little-known cities from a satellite view were selected (1 map for the practice trial, 10 for the experimental trials, see Supplementary Materials). Each map was labeled with 10 logos of the same size (1.33 cm x 1.33 cm) indicating typical stores or landmarks that can be found in a city (e.g., pharmacy, refrigerated food store, tourist office). All logos were easily identifiable as assessed by a naming agreement greater than or equal to 50 obtained in a pilot study including 20 participants who did not take part in the main experiment. In both the pilot study and the main experiment, the logos were presented in color upon a black background and none of them were repeated more than once. In the main experiment, we created two presentation orders of the maps which were counterbalanced across participants.

The Standard French speaking partner had to follow a script with 21 riddles (1 riddle for the practice map, and 2 riddles per experimental map). Therefore, the partner was a confederate. For each riddle, there was one logo that was the correct solution. The riddles contained 41 /e/-/ɛ/ minimal pairs which were different from those used in the reading task in order to capture more abstract adaptations after the social interaction. The riddle for the practice map had one /e/-/ɛ/ minimal pair. The 40 remaining pairs were distributed throughout the remaining 20 riddles and were either monosyllabic, bisyllabic, or trisyllabic words. Minimal pairs did not necessarily appear within the same riddle, but each riddle contained an equal number of /e/ and /ɛ/ target words (see Supplementary materials). The selected target words within /e/-/ɛ/ minimal pairs had a mean lexical frequency of 16.87 per million of words ($SD = 38.35$). Each riddle was 59.86 words long on average ($SD = 17.85$).

2.3. Procedure

After the participant signed the consent form, the experiment began with the setting-up of the 64-channel BioSemi system (see Figure 1). Participants then completed the reading task

followed by the other tasks (perceptual task, social interaction task, perceptual task, reading task and questionnaires) in an acoustically shielded chamber. Participants remained in the same chamber during the entire experiment.

2.3.1. Reading Task

Participants were instructed to read the sentences out loud as they appeared one at a time on the display screen (font Courier, size 18). They controlled the speed of the task and moved on from one sentence to the next with a button press. Stimuli were presented using E-Prime 2.0 software (Schneider et al., 2012). Voices were recorded with a Zoom H4n Pro audio recorder. The reading task took approximately 5 minutes.

2.3.2. Perceptual Task

The stimuli were presented binaurally using Matlab version (R2016a) (Mathworks Inc.) via Sennheiser CX 281 intra-auricular headphones at a comfortable sound level in an acoustically shielded chamber. Trials were composed of three syllables, each separated from the previous one by 600 ms of silence. On any given trial, the first two stimuli were phonemically identical, but spoken by two different female speakers. The order of the female voices was counterbalanced across the trials. The third and final stimulus, the test syllable produced by a male speaker, was either the phonemically similar to (control condition) or different from (deviant condition) the first two syllables (see Figure 1b). Similar to previous studies (Brunellière et al., 2009; Dehaene-Lambertz et al., 2000; Dufour et al. 2013), we chose to have a male speaker produces the last syllable in order to make it easy to detect the last syllable in comparison with the two other syllables produced by female speakers. In addition, this allowed us to use the same syllable in the control and deviant conditions for a given contrast. Moreover, acoustic variability between the two female speakers and acoustic variability between the two

female speakers and the male speaker forced participants to rely on more abstract phonological representations during the processing of the last syllable to make their decision. By manipulating two factors (Contrast and Condition), we obtained four experimental conditions: /e/-/ɛ/ control condition (/be/-/be/-/be/ or bɛ/-/bɛ/-/bɛ/), /e/-/ɛ/ deviant condition (/be/-/be/-/bɛ/ or bɛ/-/bɛ/-/be/), /i/-/y/ control condition (/bi/-/bi/-/bi/ or /by/-/by/-/by/) and /i/-/y/ deviant condition (/bi/-/bi/-/by/ or /by/-/by/-/bi/). Forty trials were presented per experimental condition, leading to a total of 160 experimental trials which were presented randomly. For half of forty trials, one of the two female speakers produced the first syllable, while the other female speaker produced the second syllable (and vice versa for the other half of forty trials). All experimental trials were divided into two blocks of 80 trials separated with a short break. Participants were familiarized to the task with 16 practice trials. They were instructed to indicate with a keypress whether the test syllable was the same or different from the two previous syllables as accurately and quickly as possible. The response buttons were counterbalanced across participants. Each trial started with a 500 ms white asterisk centered on a black background which remained visible on the screen during the trial until the response was given. Participants were instructed to focus their attention on the white asterisk and to avoid moving until it disappeared. An intertrial interval of 2000 ms concluded the trial. The perceptual task lasted approximately 15 minutes.

2.3.3. Social Interaction Task

The participant interacted with a Standard French speaking partner using an audio-conference device³ and intra-auricular headphones. The participant had the physical A4 size printout of the maps and was instructed to listen carefully to the description given by the

³ To ensure high-quality sound through the audio-conference device, we had checked with Standard French speakers that the /e/-/ɛ/ contrast could be differentiated as well as other phonemic productions using this device.

Standard French speaking partner about a landmark on the map. After the description, the participant selected the best logo corresponding to this description using Limesurvey (Limesurvey GmbH., 2022). Both speakers subsequently reached an agreement as to how to name the place. Once an agreement had been reached, the participant typed out the answer on Limesurvey. The social interaction task took approximately 30 minutes.

2.3.4. Perceptual and Reading Tasks and questionnaires

After the social interaction task, the participant performed again the perceptual task followed by the reading task and they then completed two questionnaires. We assessed their empathy ability by using the 28-item French version of the Interpersonal Reactivity Index questionnaire (IRI, Braun et al., 2015). To make sure that the partner was judged as being agreeable during social interaction, participants filled out a partner likability questionnaire by rating how pleasant (“*sympathique*”), friendly (“*amicale*”), social (“*social*”), and relaxed (“*détendu*”) they found their partner on a 5-point Likert scale (for a similar approach, see Schweitzer & Lewandowski, 2014). At the end of the questionnaires, participants answered additional questions in order to determine if they had identified which region their partner came from and if their partner was a confederate. Most participants (84.6%) were not capable of guessing the region their partner came from and they spontaneously reported not being aware of the presence of a confederate. In total, the experiment lasted approximately 120 minutes.

2.4. EEG recording system and preprocessing

The continuous EEG signal was recorded (sampling rate: 1024 Hz) using a 64-channel BioSemi ActiveTwo A/D box system (BioSemi Instrumentation, Amsterdam, The Netherlands) and sintered Ag/AgCl active electrodes organized in the 10/5 system. Four external flat-type electrodes were placed near both mastoids and the right eye to control for eye blinks and

mandibular movements. Individual electrodes were adjusted to a stable offset of < 20 mV. As in Brunellière et al. (2009), the EEG epochs, starting at 100 ms before test syllables and ending 500 ms after these syllables were averaged, from each experimental condition (/e-/ε/ control condition, /e-/ε/ deviant condition, /i-/y/ control condition and /i-/y/ deviant condition), for each participant and for each electrode. The EEG data was filtered offline by a Butterworth bandpass (0.1 Hz for the highpass filter, 30 Hz for the lowpass filter) and when necessary, a 50 Hz notch filter. Each EEG epoch was corrected to a 100-ms baseline placed before the test syllable onset. To eliminate eye blinks and other artifacts, epochs were removed under a rejection criterion of ± 100 μ V at any channel within the period of epochs. All participants had a number of accepted trials superior to 30 for each experimental condition. The total number of accepted epochs was equivalent across the experimental conditions before and after the social interaction (means: before social interaction, /e-/ε/ control condition, 38.6, /e-/ε/ deviant condition, 38.6, /i-/y/ control condition, 38.9, /i-/y/ deviant condition, 39; after social interaction, /e-/ε/ control condition: 38.1, /e-/ε/ deviant condition: 38.5 /i-/y/ control condition: 38.9, /i-/y/ deviant condition: 39). The EEG signal was transformed offline using the average reference. Data from bad channels was interpolated for each participant (Perrin et al., 1987) as long as none of the surrounding channels were bad. No more than 3 channels were interpolated (mean: 1.25).

2.5. Data Analysis

2.5.1. Reading Task

We measured the average Euclidean distance in the F1-F2 plane for the /e-/ε/ and /i-/y/ contrasts. Based on our hypotheses, we performed student *t*-tests on the average Euclidean distance in Barks to assess the effect of contrast before and after the social interaction. In

addition, we used Pillai score (Hay et al. 2006; Nycz & Hall-Lew, 2013) which is a summary statistic of multivariate analyses of variance (MANOVA) including F1 and F2 as dependent variables in order to examine the spectral overlap between the vowels within each phonemic contrast (/e/-/ɛ/ and /i/-/y/) across the F1-F2 plane. A Pillai score close to 0 suggests a strong spectral overlap corresponding to merged categories.

2.5.2. *Perceptual task*

Behavioral data

We performed one ANOVA on mean accuracy (correct responses) by participant with Contrast (/e/-/ɛ/ vs. /i/-/y/) and Condition (control vs. deviant) as factors before and after the social interaction. Based on our previous hypotheses, planned comparisons were conducted between the two conditions in each contrast and between the two contrasts in the deviant condition. Due to participants' poor performance leading to not having enough accurate trials in the /e/-/ɛ/ deviant condition both before and after the social interaction, we were not able to perform analyses on reaction times.

EEG data

Based on previous studies (Brunellière et al., 2009; Dufour et al., 2013), four time windows centered around the peak amplitude of the components of interest were selected: 140 ms - 180 ms (N100), 220 ms - 260 ms (P200), 335 ms - 375 ms (MMN), and 445 ms - 485 ms (late MMN time window). We performed an ANOVA on mean amplitude by participant with the following factors over each time window: Contrast (/e/-/ɛ/ vs. /i/-/y/), Condition (control vs. deviant) and Site (frontocentral, left temporal, right temporal, centroparietal, left posterior, and right posterior). As in Brunellière et al. (2009), we used 6 sites of 6 electrodes to explore specific topographical effects. The scalp surface was divided as follows: frontocentral (F1, F2, Fz, FC1,

FCz, and FC2), left temporal (AF7, F7, F5, FT7, FC5 and AF3), right temporal (AF8, F6, F8, FT8, FC6 and AF4), centroparietal (CP1, CPz, CP2, Pz, C1, and C2), left posterior (CP5, CP3, P1, PO3, P5 and P3), and right posterior (CP6, CP4, P2, PO4, P6 and P4). When there was more than one degree of freedom in the numerator, the Greenhouse–Geisser correction was applied (Greenhouse & Geisser, 1959) and the corrected *p*-values are reported below in the result section. Based on our previous hypotheses, planned comparisons over MMN and MMN late time windows were conducted to test significant interactions between the factors, Contrast and Condition, main effects of Condition and effects of Condition over particular sites (left temporal, right temporal, frontocentral, centroparietal, left posterior, right posterior). Without a priori hypotheses, post-hoc Tukey comparisons were performed when significant effects were found. We reported only significant effects.

2.5.3 Social interaction task and questionnaires

We computed the mean and the standard deviation of accuracy in the selection of the correct logo per participant during the social interaction task. Additionally, the mean and the standard deviation of the two questionnaires (partner likability questionnaire and empathy questionnaire with Interpersonal Reactivity Index) were extracted. We report these findings below before presenting those obtained in the reading and perceptual tasks.

3. Results

3.1 Social interaction task and questionnaires

Social interaction task

Participants had no difficulty solving the riddles described by the Standard French speaking partner and reaching mutual understanding during the social interaction task. Indeed,

accuracy in logo selection was 98.44% ($SD = 2.38\%$). Thirteen participants answered all riddles correctly, with the remaining participants failing to solve only one of the riddles.

Questionnaires

Participants' ratings on the 4-item partner likability questionnaire indicated that they judged their partner as being very agreeable. The values on the 5-point Likert scale were the following: pleasantness ($M = 4.95$, $SD = 0.23$), friendliness ($M = 4.63$, $SD = 0.60$), socialness ($M = 4.74$, $SD = 0.56$), and relaxedness ($M = 4.47$, $SD = 0.70$). The evaluation of the participants' empathy across the 4 distinct sub-dimensions of the Interpersonal Reactivity Index produced the following values (the maximum value for each sub-dimension being 28): fantasy ($M = 18.68$, $SD = 4.80$), perspective-taking ($M = 19.53$, $SD = 3.32$), empathetic concern ($M = 21.16$, $SD = 4.11$), and personal distress ($M = 13.42$, $SD = 6.35$). These values were descriptively very similar to those reported in the study by Braun et al. (2015) with a large population.

3.2 Reading task

Before the social interaction. The statistical analysis on the Euclidean distance showed a significant difference between the two phonemic contrasts (/e/-/ε/ versus /i/-/y/) before the social interaction, ($t(18) = -10.24$, $p = 4.63 \times 10^{-9}$). As expected, the Euclidean distance for the /i/-/y/ contrast (in Barks, $M = 1.46$, $SD = 0.29$) was higher than that for the /e/-/ε/ contrast (in Barks, $M = 0.77$, $SD = 0.18$). Moreover, there was a significant difference in F1-F2 spectral distribution for the /i/-/y/ control contrast ($F(1,337) = 38.64$, $p = 7.86 \times 10^{-16}$, Pillai score: 0.187). On the contrary, there was no significant difference in spectral overlap for the /e/-/ε/ contrast ($F(1,331) = 0.34$, $p = .71$, Pillai score: 0.0021). As seen in Figure 2a and 2c, the F1-F2 spectral distribution between the /i/ and /y/ vowels was distinct, while there was a full overlap between the /e/ and /ε/ vowels.

< Insert Figure 2 >

After the social interaction. The analysis on the Euclidean distance again revealed a significant difference found between the two contrasts ($t(18) = -6.72, p = 2.68 \times 10^{-6}$, in Barks, /i/-/y/ contrast, $M = 1.37, SD = 0.27$; e/-/ε/ contrast, $M = 0.83, SD = 0.26$). There was also a significant difference in F1-F2 spectral distribution for the /i/-/y/ control contrast ($F(1,338) = 34.61, p = 2.14 \times 10^{-14}$, Pillai score 0.170). The lack in significant difference between the spectral distributions between the /e/ and /ε/ vowels was maintained after the social interaction ($F(1,335) = 0.46, p = .63$, Pillai score: 0.0028). The spectral distribution between the /i/ and /y/ vowels was distinct, while there still was a full overlap between the /e/ and /ε/ vowels (see Figure 2b and Figure 2d).

3.3 Perceptual task

Behavioral data

Before the social interaction. The ANOVA on the accuracy before the social interaction showed significant main effects for Contrast ($F(1,18) = 550.75, p = 6 \times 10^{-15}$) and Condition ($F(1,18) = 467.63, p = 2.49 \times 10^{-14}$), and a significant Contrast x Condition interaction ($F(1,18) = 421.12, p = 6.17 \times 10^{-14}$). Planned comparisons on the accuracy were conducted between the two conditions for each contrast and between the two contrasts in the deviant condition. As seen in Figure 3a, participants' accuracy for the /i/-/y/ contrast was above 0.9 in the deviant condition and did not differ between the control and deviant conditions, $F(1,18) = 2.8, p = .11$. As expected, participants' accuracy was below chance level in the deviant condition for the /e/-/ε/ contrast (see Figure 3a). Their accuracy in the deviant condition significantly differed from that in the control condition for the /e/-/ε/ contrast, $F(1,18) = 490.35, p = 1.65 \times 10^{-14}$. In the deviant

condition, their accuracy for the /i/-/y/ contrast was significantly higher than that for the /e/-/ε/ contrast, $F(1,18) = 535.83, p = 7.66 \times 10^{-15}$.

< Insert Figure 3 >

After the social interaction. The statistical analysis also revealed significant main effects for Contrast ($F(1,18) = 104.81, p = 6.21 \times 10^{-9}$), and Condition ($F(1,18) = 96.47, p = 1.18 \times 10^{-8}$), and a significant Contrast x Condition interaction ($F(1,18) = 274.60, p = 2.41 \times 10^{-9}$). Participants' accuracy for the /i/-/y/ contrast was again above 0.9 in both the control and deviant conditions (see Figure 3b). Their accuracy for the /i/-/y/ contrast did not significantly differ between the control and deviant conditions, $F(1,18) = 0.2, p = .66$. Contrary to this pattern, the accuracy in the deviant condition was significantly different from that in the control condition for the /e/-/ε/ contrast, $F(1,18) = 189.77, p = 5.32 \times 10^{-11}$. Participants' accuracy was still below chance level in the deviant condition for the /e/-/ε/ contrast (see Figure 3b) and their accuracy for the /i/-/y/ contrast was significantly higher than that for the /e/-/ε/ contrast in the deviant condition, $F(1,18) = 198.54, p = 3.65 \times 10^{-11}$.

EEG data

The grand-average ERP waveforms per experimental condition are shown in Figures 4 and 5, respectively for before and after the social interaction. A first negative response peaked after 100 ms (N100) and was then followed by a positive peak after 200 ms (P200). Then, a larger negative mismatch response was elicited after the deviant condition for /i/-/y/ contrast. Based on visual inspection, it seemed that a larger negative mismatch response was also elicited after the deviant condition for the /e/-/ε/ contrast after the social interaction. A summary of average amplitudes over each time window and site per contrast and condition as well as a

summary of all statistical results over each time window are provided in Table 1 and Table 2, respectively.

< Insert Figures 4 and 5 >

< Insert Table 1 >

< Insert Table 2 >

Before the social interaction

N100 time window. There were main effects of Contrast, $F(1,18) = 4.36, p = .05$ and of Site, $F(5,90) = 5.74, p = 9.12 \times 10^{-3}$. The amplitude of N100 was higher for the /i/-/y/ contrast than that for the /e/-/ɛ/ contrast and the centroparietal site had more negative values than the left temporal site ($p = 1.91 \times 10^{-4}$), the right temporal site ($p = 3.48 \times 10^{-3}$) and the left posterior site ($p = 4.4 \times 10^{-2}$), as usually found for the N100 topography.

P200 time window. No significant effect was found.

MMN time window. The analysis revealed a main effect of Condition, $F(1,18) = 4.93, p = 3.94 \times 10^{-2}$, but we found a significant interaction between Contrast, Condition, and Site, $F(1,18) = 3.08, p = .05$. In line with our hypotheses, we conducted planned comparisons to test significant interactions between the factors, Contrast and Condition and effects of Condition over particular sites (left temporal, right temporal, frontocentral, centroparietal, left posterior, right posterior). The significant interaction between Contrast, Condition, and Site was due to the fact that there was only an effect of Condition for the /i/-/y/ contrast over the left temporal site, $F(1,18) = 4.87, p = 4.06 \times 10^{-2}$ and over the left posterior site, $F(1,18) = 7.56, p = 1.32 \times 10^{-2}$. For the /i/-/y/ contrast, there were more negative values over the left temporal site in the deviant condition than in the control condition, while more positive values over left posterior site were shown in the deviant condition than in the control condition (see Figure 4). A significant interaction

between Contrast and Condition over the left posterior site was also found, $F(1,18) = 7.25$, $p = 1.49^{e-2}$. The main effect of Site ($F(5,90) = 36.59$, $p = 7.17^{e-9}$) revealed the classic topography of the MMN component, as usually reported with more negative values over the more frontal sites (left temporal, right temporal and frontocentral sites, $ps < .001$).

Late MMN time window. The analysis showed a main effect of Site, $F(5,90) = 29.71$, $p = 6.73^{e-9}$, showing that the classic topography of the MMN component with more negative values over the more frontal sites (left temporal, right temporal and frontocentral sites, $ps < .001$). In line with our hypotheses, we conducted planned comparisons to test significant interactions between the factors, Contrast and Condition and effects of Condition over particular sites (left temporal, right temporal, frontocentral, centroparietal, left posterior, right posterior). We found a significant interaction between Contrast and Condition over the left posterior site, $F(1,18) = 8.48$, $p = 9.3^{e-3}$. This was explained by an effect of Condition only for the /i/-/y/ contrast, $F(1,18) = 9.78$, $p = 5.82^{e-3}$ over the left posterior site. As seen in Figure 4, there were more positive values over the left posterior site in the deviant condition than in the control condition for the /i/-/y/ contrast.

After the social interaction

N100 time window. There was a main effect of Contrast, $F(1,18) = 5.81$, $p = 2.68^{e-2}$. The amplitude of N100 was higher for the /i/-/y/ contrast than that for the /e/-/ε/ contrast.

P200 time window. Surprisingly, there were main effects of Contrast, $F(1,18) = 5.99$, $p = 2.49^{e-2}$ and of Condition, $F(1,18) = 4.41$, $p = .05$ and a significant interaction between Contrast and Condition, $F(1,18) = 6.52$, $p = 2^{e-2}$. As a significant interaction between Contrast and Condition was found, post-hoc analyses were conducted and the main effects were not

described. The significant interaction between Contrast and Condition was explained by an effect of Condition only for the /i/-/y/ contrast ($p = 5.47 \times 10^{-3}$). The amplitude of the P200 was reduced in the deviant condition with respect to the control condition. There were more negative values in the deviant condition.

MMN time window. The analysis revealed a main effect of Site, $F(5,90) = 19.23$, $p = 4.29 \times 10^{-5}$, showing the classic topography of the MMN component with more negative values over the more frontal sites ($ps < .001$). In line with our hypotheses, we conducted planned comparisons to test main effects of Condition over particular sites (left temporal, right temporal, frontocentral, centroparietal, left posterior, right posterior). A main effect of Condition was found over the left temporal site, $F(1,18) = 4.15$, $p = .05$. As seen in Figure 5, the amplitude of the MMN over the left temporal site was stronger in the deviant condition than in the control condition for the two contrasts.

Late MMN time window. The analysis revealed a main effect of Site, $F(5,90) = 12.87$, $p = 5.34 \times 10^{-4}$, a significant interaction between Contrast and Site, $F(5,90) = 3.49$, $p = 2.87 \times 10^{-2}$ and a significant interaction between Condition and Site, $F(5,90) = 5.10$, $p = 2.73 \times 10^{-3}$. The main effect of Site showed the classic topography of the MMN component with more negative values over the more frontal sites ($ps < .001$). Post-hoc analyses over the significant interaction between Contrast and Site did not reveal any significant difference between the two contrasts over one topographical site. In line with our hypotheses, the significant interaction between Condition and Site was explained by a main effect of Condition over the left ($F(1,18) = 4.7$, $p = 4.37 \times 10^{-2}$) and right ($F(1,18) = 9.14$, $p = 7.31 \times 10^{-3}$) temporal sites and the left posterior site ($F(1,18) = 10.12$, $p = 5.18 \times 10^{-3}$). For the two contrasts, the MMN component over the left and right temporal sites had more negative values in the deviant condition than in the control condition, while over the left

posterior site it had more positive values between deviant and control condition (see Figure 5). Since we found a main effect of Condition over a particular site or several sites during the MMN time window or the late MMN time window after the social interaction, supplementary analyses focusing on the /e/-/ɛ/ contrast which include empathy and partner likeability as factors were conducted over these time windows. The findings of the analyses are reported in the Supplementary Materials section.

4. Discussion

In the present study, we investigated whether a neuronal discrimination of the /e/-/ɛ/ phonemic merger in Northern French speakers can occur after a social interaction with a partner coming from another region and producing the /e/-/ɛ/ contrast. We developed a well-controlled interactive setting through which /e/-/ɛ/ minimal word pairs inserted into pre-crafted riddles were produced by a Standard French speaker. Northern French speakers who did not produce and perceive this contrast interacted with their partner to find some places on a map and reach an agreement to name places. This experimental setting was designed to maximize Northern French speakers' adaptations by making them share a common goal with their partner in a pleasant conversational situation. Nonetheless, Northern French speakers had poor performance in differentiating between the /e/ and /ɛ/ vowels after the social interaction, which highlighted their difficulty to consciously distinguish between the two vowels at the decisional stage. Additionally, acoustic data in the reading task revealed that Northern French speakers still did not distinctly produce the /e/-/ɛ/ contrast. However, in contrast to acoustic data in the reading task or behavioral data in the perceptual task, there was clear evidence for neuronal discrimination of the /e/-/ɛ/ phonemic merger in Northern French speakers after the social

interaction. Similar to the /i/-/y/ stable phonemic contrast, MMN responses were elicited by the /e/-/ɛ/ phonemic merger after the social interaction. In light of the previous literature, the implications of these findings are discussed below.

The present study is the first to provide evidence that social interaction can promote rapid neuronal adaptations in the perception of a new phonemic contrast which is subject to variations within the native language. First of all, this finding is in line with a large number of studies showing brain plasticity for learning new speech features and phonemic contrasts (Atienza and Cantero, 2001; Cheour et al., 2002; Kraus et al., 1995; Menning et al., 2000; Näätänen et al., 1993; Partanen et al. 2013, Peltola et al., 2005; Shestakova et al., 2003; Tamminen et al., 2015; Tang et al., 2020; Tremblay et al., 1997, 1998; Winkler et al., 1999; Zhang et al., 2009). Similar to these studies, we explored the MMN response, which is known to reflect the memory traces of native speech sounds (Näätänen et al., 2007; Shtyrov & Pulvermüller, 2007). However, the majority of previous ERP studies have not been conducted on the learning of phonemic mergers with the exception of the study by Tang et al. (2020). Contrary to this study, we did not attempt to explore the impact of a long-period training with a 7-day training by using orthographical forms to rebuild the memory traces of phonemic representations. More precisely, we questioned whether exposure to partner's productions in a collaborative dialogue setting could lead to rapid changes in the memory traces of phonemic representations. It should be noted that we cannot exclude that merely listening to phoneme contrasts in the riddles might have played a role on the memory traces of phonemic representations. However, the significant interactions found with the partner likability ratings support the idea that the rapid changes in the memory traces of phonemic representations are caused by the interactive situation between the two interlocutors rather than by the mere exposure to partner's productions. In terms of neural substrates, since

changes in memory traces have been attributed to variations in synaptic connectivity between neuronal assemblies (Froemke et al., 2007; Libet et al., 1975; Yu et al., 2004), newly formed neuronal assemblies distinctly representing the /e/ and /ɛ/ vowels therefore seemed to be established after a social interaction.

One intriguing finding is the divergence between this rapid neuronal discrimination of the /e/-/ɛ/ phonemic merger evidenced by the MMN responses and the behavioral findings in the perceptual task. Although there was clear evidence in support of a neuronal discrimination of the /e/-/ɛ/ phonemic merger after the social interaction, poor performance was found in the attentive perception of this contrast. In prior literature (Tang et al., 2020), an absence of correlation between the behavioral performance with accuracy and the MMN response was also observed. Some previous ERP studies focusing on detection of features also highlighted a dissociation between the MMN response and behavioral data (Calcus et al., 2015; Colin et al., 2009; Timm et al., 2011). The MMN response was influenced by the nature of features which were manipulated in these previous studies (Calcus et al., 2015; Colin et al., 2009; Timm et al., 2011), whereas the behavioral measures revealed that participants failed to consciously detect these features. This dissociation highlights the core processing reflected by the MMN response which is a pre-attentive process associated with the activation of memory traces and which differs from the accuracy measures such as during a forced choice task mainly resting on conscious decisional processes. The same explanation was provided in Dufour et al. (2013). Conscious decisional processes can especially be based on metalinguistic knowledge which includes the use of a variety of linguistic and contextual information.

Nonetheless, Tang et al.'s study (2020) revealed discrimination towards the /n/-/l/ merger on both the perceptual and neural levels after a 7-day training during which participants received

positive feedback and were exposed to syllables produced by female and male speakers. In addition, a recent preprint (Zhou et al., preprint) using Chinese characters varying according to the initial consonant /l/ or /n/ reported an improvement of behavioral discrimination and a visual MMN response become similar to that of the unmerged speakers after a 5-day training. Converging findings on both behavioral and neural levels for perception seemed to be observed when an intensive training was administered to the participants. In contrast, our findings suggest that the changes observed immediately after the social interaction were limited to neuronal memory traces of phonemic representations. This raises the question of whether these changes last over time and whether the duration of a social interaction, such as during an intensive training of several days, can give rise to changes on both behavioral and neural levels for perception or even for production. Future studies are needed to understand if changes in neuronal memory traces of phonemic representations which occur after a social interaction persist over time and whether longer social interactions (e.g. more social interaction map sessions over several days) have a strong impact at both behavioral and neural levels for perception and production. Beyond the question of social interaction length, cumulative effects of exposure to the /e/-/ɛ/ contrast during the three tasks including the reading, perceptual and social interaction tasks in the current study might have facilitated the emergence of the changes observed in neuronal memory traces of phonemic representations. Even though participants did not distinctly produce or perceive the /e/-/ɛ/ contrast before the social interaction, we cannot exclude the impact of cumulative effects related to the exposure to this contrast after the multiple tasks. However, using a within-participant design offered the possibility to examine the participants' abilities to perceive and produce the /e/-/ɛ/ contrast before and after the social interaction and to avoid group comparison issues due to inter-individual differences. Another interesting aspect

from the discrepancy between Tang et al.'s study (2020) and Zhou et al. (preprint) with respect to our study concerns the impact of the feedback on syllable discrimination during the training and the nature of perceptual task involving memory retention of stimuli in order to make a same-different judgment. More studies are needed to examine the impact of the feedback during the training and that of various perceptual tasks associated with memory retention of stimuli.

Moreover, the results found in the reading task lead us to reconsider the relationship of speech perception and production and the level of linguistic information engaged during this task. Although some studies claim functional links between speech perception and production (Fadiga et al., 2002; Watkins & Paus, 2004), our findings indicate that changes in memory traces of phonemic representations evidenced by the MMN response did not interact with the speech production system. This may be due to the fact that the sources of changes in the MMN response for learning new contrasts were usually localized in the left hemisphere of the superior temporal region (Zhang et al., 2009). In line with a dissociation between speech perception and production, one behavioral study examining individuals' ability to perceive a speech sound and the distinctness between contrasting sounds in production revealed no relationship between speech perception and production (Cheng et al., 2021). Another explanation for our findings in the reading task may come from the level of linguistic information engaged during the reading task. Participants were asked to read sentences, leading to the activation of different levels of representations, including their knowledge of word forms. In addition, as we explored the use of abstract representations, minimal word pairs embedded in the sentences were different from those provided during the social interaction. Future studies are needed to determine whether such interactive setting can produce changes in the activation of words differing in the phonemic merger when the minimal word pairs produced by the partner are tested. The development of

such interactive settings thus opens new perspectives on the debate of the abstractionist and exemplar-based models in the lexicon (Goldinger, 1998; Marslen-Wilson, 1984; Morton, 1979; Norris et al., 2003) and on the nature of the relationship between speech perception and production.

It is also interesting to understand by which neural mechanisms social interaction can lead to neural discrimination of a vowel merger. One possibility may be related to social abilities, such as empathy which involves taking a partner's perspective into account. In the present study, we found no link between the neural discrimination of the /e/-/ɛ/ phonemic merger and the four subdimensions of the IRI, which is the most used questionnaire to assess empathy (Davis, 1980). Another possibility would rest on the phenomenon of neural coupling between individuals involved in an interaction (e.g., Bevilacqua et al., 2019; Chen et al., 2022; Stephens et al., 2010). For instance, neural coupling between speakers in regions involved in language production and perception during social interaction is correlated with a better comprehension of the interaction (Stephens et al., 2010). In other words, one unresolved question is whether the activity aligned between two brains may be the neural substrate that triggers changes in memory traces of phonemic representations. Further studies should be conducted to understand whether the neural discrimination of a vowel merger following a social interaction depends on the strength of neural coupling between speakers during this interaction. More studies are also needed to have a greater generalizability of the results across phonemic contrasts in a variety of languages and with various dyadic situations taking the talkers' gender into account.

5. Conclusions

Taken together, the results of the present study suggest that interacting with a partner producing a new phonemic contrast within a native language gives rise to rapid adaptations of phonemic representations. The neural discrimination of the /e/-/ɛ/ phonemic merger was clearly evidenced by the MMN responses after the social interaction. Although newly formed neuronal assemblies distinctly representing the /e/ and /ɛ/ vowels seemed to be established after a social interaction, conscious decisional processes during the perception of the phonemic merger and production mechanisms involving sentences containing minimal pairs with the phonemic merger were not affected by the social interaction. Our results provide the basis for exploring the neural adaptations at phonological level after social interactions.

Supplementary Materials

The ERP findings on supplementary analyses including empathy and partner likability questionnaires are below described. As we found a main effect of Condition over the left temporal site during the MMN time window after the social interaction, we conducted a supplementary analysis to test if the effect of Condition on the left temporal site was influenced by the participant's degree of empathy for the /e/-/ɛ/ contrast. No significant interaction was found between Condition and Empathy (Fantasy: $F(1,14) = 0.98, p = .34$; Empathetic Concern: $F(1,14) = 0.24, p = .63$; Perspective Taking; $F(1,14) = 0.06, p = .81$; Personal Distress: $F(1,14) = 0.06, p = .81$). A supplementary analysis was also conducted over the late MMN time window to test if the effects of Condition for the /e/-/ɛ/ contrast found after the social interaction were influenced by the degree of empathy. It showed no significant interaction between Condition and Empathy (Fantasy: $F(1,14) = 0.45, p = .51$; Empathetic Concern: $F(1,14) = 0.28, p = .61$; Perspective Taking; $F(1,14) = 0.26, p = .51$; Personal Distress: $F(1,14) = 1.69, p = .21$). Even though three different topographical sites (left posterior, left temporal, and right temporal) were included in the analysis of this time window, no significant interaction between Condition, Site and Empathy was found (Fantasy: $F(2,28) = 0.93, p = .41$; Empathetic Concern: $F(2,28) = 0.30, p = .74$; Perspective Taking; $F(2,28) = 0.32, p = .73$; Personal Distress: $F(2,28) = 2.16, p = .13$).

In addition to the participant's degree of empathy, we performed supplementary analyses to investigate a potential impact of the partner likability ratings. Over the MMN time window for the left temporal site, no significant interaction was found between Condition and Partner likability in terms of Friendliness ($F(1,14) = 0.008, p = .93$), Socialness, ($F(1,14) = 0.83, p = .38$), and Relaxedness ($F(1,14) = 0.002, p = .97$). However, there was a significant interaction between Condition and Pleasantness ($F(1,14) = 7.91, p = 1.38 \times 10^{-2}$) over this time window.

Participants who rated their partner as being more pleasant were those who exhibited the strongest effect of the condition. Then, over the late MMN time window, a significant interaction between Condition, Site and Relaxedness was shown ($F(2,28) = 4.54, p = 2.41^{e-2}$). This interaction was explained by a significant interaction between Condition and Relaxedness over the left posterior site ($F(1,17) = 9.08, p = 7.83^{e-3}$). Participants who rated their partner as being more relaxed were those who exhibited the strongest effect of the condition over this site. Neither a significant interaction between Condition and Relaxedness over the left ($F(1,17) = 4.17, p = 6.59^{e-2}$) and right temporal site ($F(1,17) = 1.12, p = .30$) nor significant interactions between Condition, Sites and another subscale of partner likability were found (Pleasantness $F(2,28) = 0.42, p = .64$; Friendliness, $F(2,28) = 1.6, p = .22$; Socialness, $F(2,28) = 1, p = .36$). The materials used in the reading and social interaction tasks can be found at the following url, <https://doi.org/10.5281/zenodo.10893922>. The preprocessed data of reading and perceptual tasks are also online available at the same url.

Declaration of interest statement

All authors declare that they have no conflicts of interest. No generative AI or AI-assisted technologies were used during the writing process of this manuscript.

Acknowledgements

This work was supported by the French National Research Agency [ANR grant number ANR-19-CE28-0006]. The EEG analysis was performed using free academic software Cartool developed at the Functional Brain Mapping Lab in Geneva, Switzerland, which is supported by the Center for Biomedical Imaging of Geneva-Lausanne. We are very grateful to H el ene Braun, L eonore Catteau, Pauline Justin, Chlo e Leterme, Gwendoline Nobr e, and L ea Vincent for their help in participating as a Standard French partner during social interaction and/or in

preprocessing the productions of participants in the reading task. We thank Charlotte Stinkeste for her help in running the pilot study and implementing the social interaction task. We are also grateful to two anonymous reviewers for their help during the peer review process.

Tables and Figures

Figure 1. (a) Diagram of experimental procedure. (b) Stimuli and Conditions used in the perceptual task. The last syllable on which ERP analyses were performed was shown in bold, (c) Example of riddles provided in the social interaction task. Training riddle is shown with /e/-/ɛ/ phonemic contrast manipulation: *thé* /te/ - *taies* /tɛ/. English translation of riddle and participant's response are as follows. Standard French Partner: "You can drink both tea and cocktails here and it's often beautifully decorated with armchairs and cushions with colorful pillowcases, do you see what I am referring to?" Participant: "I think I've found it... a café?"

a Experimental procedure



b Stimuli of the perceptual task

Control Condition	Deviant Condition
/be/ - /be/ - /be/	/be/ - /be/ - /bɛ/
/bɛ/ - /bɛ/ - /bɛ/	/bɛ/ - /bɛ/ - /be/
/bi/ - /bi/ - /bi/	/bi/ - /bi/ - /by/
/by/ - /by/ - /by/	/by/ - /by/ - /bi/

c An example riddle of the social interaction task

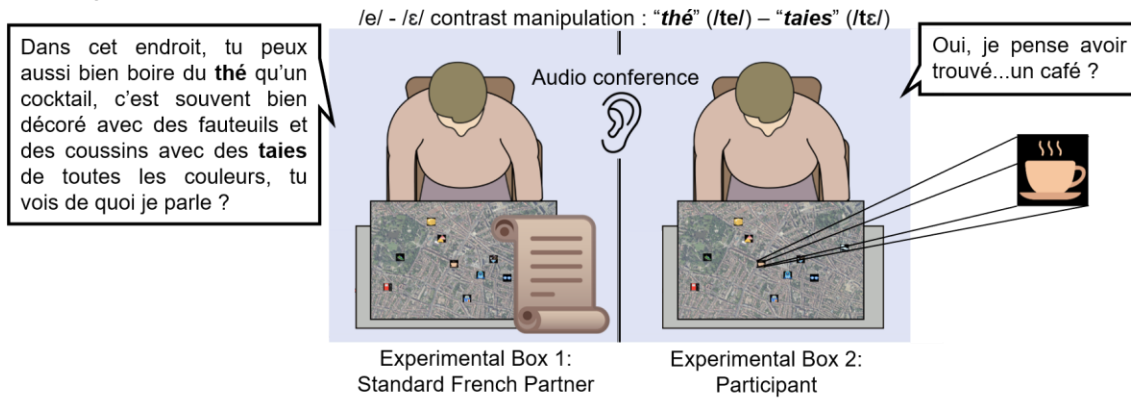


Figure 2. Spectral distribution around centroid of vowels /e/, /ɛ/, /i/ and /y/ produced in the reading task across the F1-F2 plane before (a, c) and after (b, d) the social interaction.

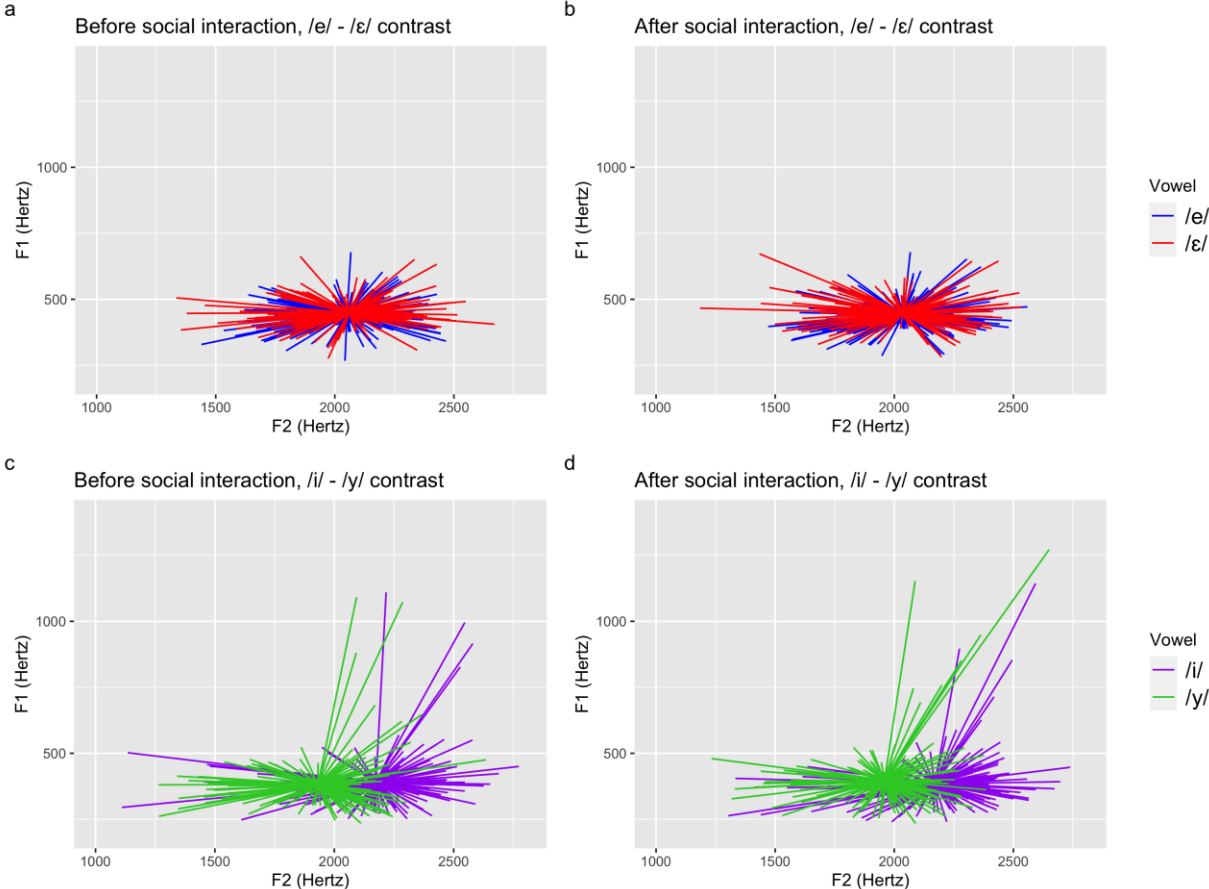


Figure 3. Perceptual task accuracy with error bars shown for before **(a)** and after **(b)** the social interaction. Contrasts (/e/-/ɛ/ vs. /i/-/y/) and Condition (control vs. deviant) are indicated.

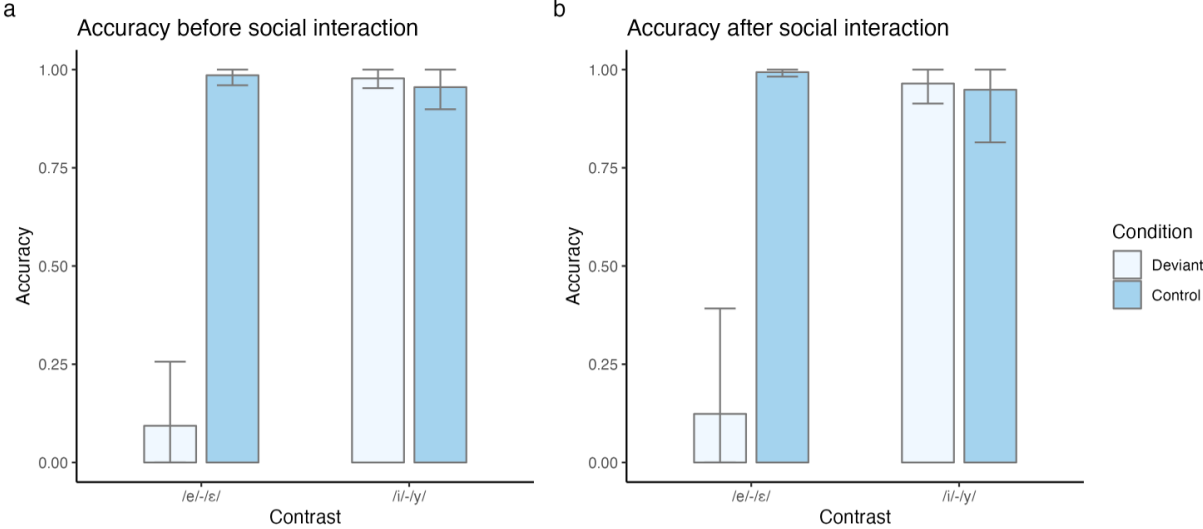


Figure 4. Grand-average waveforms for the control and deviant conditions in each phonemic contrast (*/e/-/ε/* and */i/-/y/*) over each topographical site before the social interaction. Red rectangles denote significant effects between control and deviant conditions over specific topographical sites. LT: Left temporal, RT: Right temporal, FC: Frontocentral, CP: Centroparietal, LP: Left posterior, RP: Right posterior. Last syllable started at 0 ms.

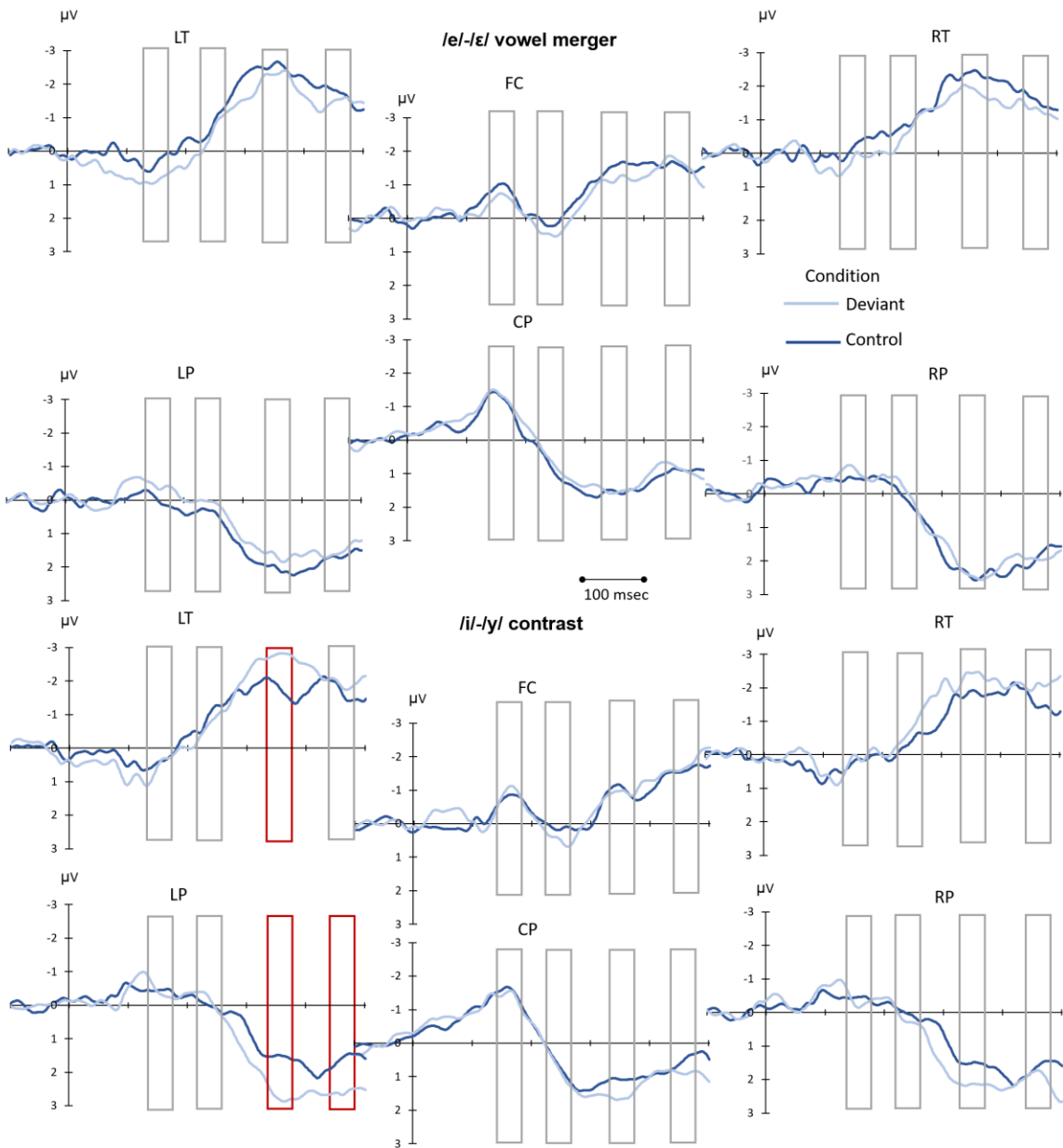


Figure 5. Grand-average waveforms for the control and deviant conditions in each phonemic contrast (/e/-/ɛ/ and /i/-/y/) over each topographical site after the social interaction. Red rectangles denote significant effects between control and deviant conditions over specific topographical sites. LT: Left temporal, RT: Right temporal, FC: Frontocentral, CP: Centroparietal, LP: Left posterior, RP: Right posterior. Last syllable started at 0 ms.

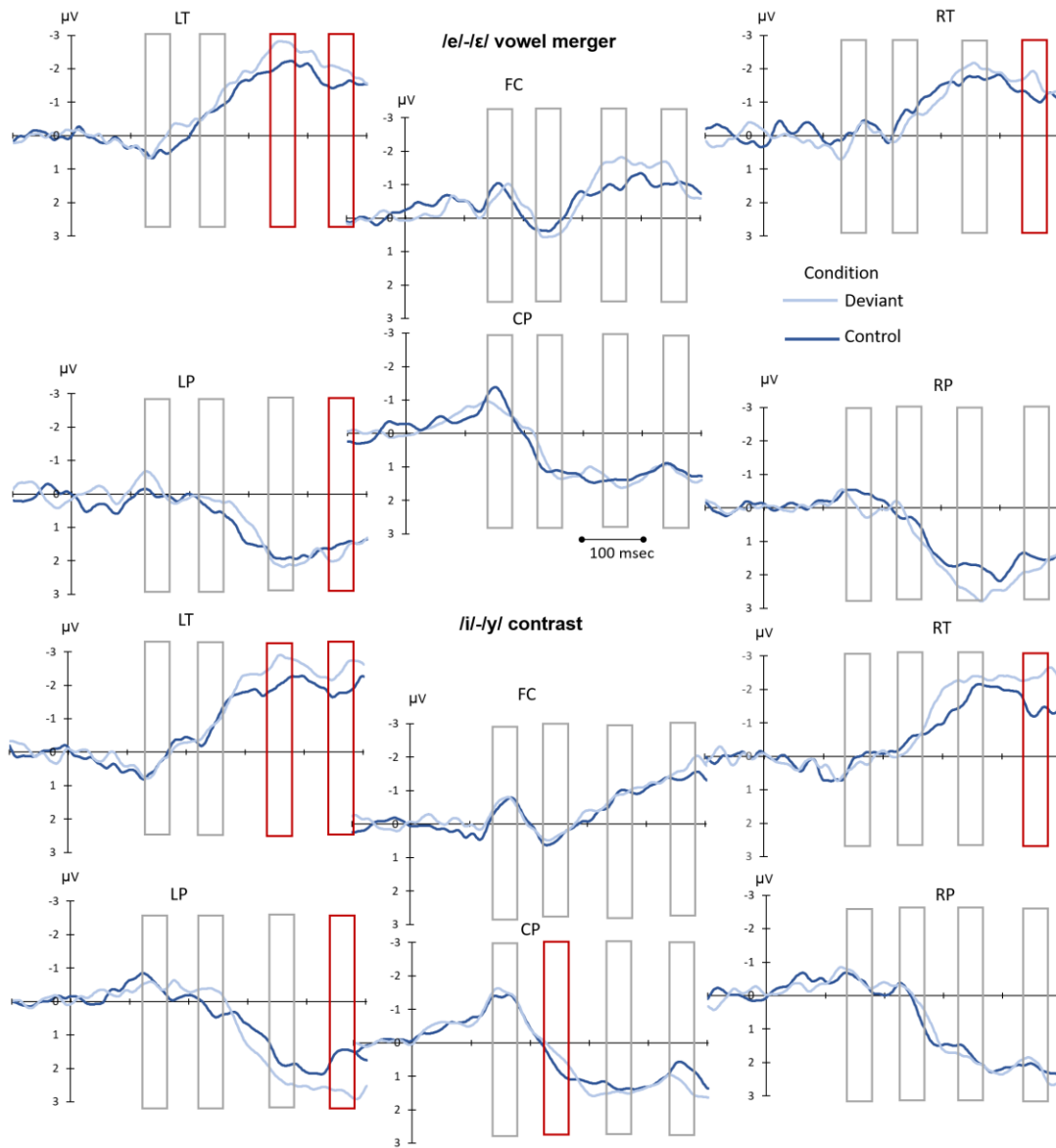


Table 1. Summary of average amplitudes (μV) over each time window and site per contrast and condition. The amplitudes provided are rounded to two decimal places. Sites abbreviated as follows: Left Temporal (LT), Right Temporal (RT), Frontocentral (FC), Centroparietal (CP), Left Posterior (LP), Right Posterior (RP).

N100 time window, Before Social Interaction							
Contrast	Condition	LT	RT	FC	CP	LP	RP
/e/-/ε/	Control	0.21	-0.25	-0.91	-1.2	0.04	-0.46
/e/-/ε/	Deviant	0.76	0.16	-0.42	-1.27	-0.40	-0.6
/i/-/y/	Control	0.31	0.16	-0.78	-1.54	-0.42	-0.5
/i/-/y/	Deviant	0.36	0.06	-0.93	-1.42	-0.31	-0.28
P200 time window, Before Social Interaction							
Contrast	Condition	LT	RT	FC	CP	LP	RP
/e/-/ε/	Control	-0.69	-0.8	0.14	0.76	0.34	0.23
/e/-/ε/	Deviant	-0.47	-0.5	0.67	0.6	0.05	0
/i/-/y/	Control	-1.05	-0.32	0.11	0.55	0.06	0.34
/i/-/y/	Deviant	-0.82	-0.63	0.46	0.64	-0.04	0.26
MMN time window, Before Social Interaction							
Contrast	Condition	LT	RT	FC	CP	LP	RP
/e/-/ε/	Control	-2.52	-2.37	-1.6	1.53	2.04	2.45
/e/-/ε/	Deviant	-2.32	-1.9	-0.97	1.54	1.69	2.48
/i/-/y/	Control	-1.73	-1.82	-0.97	1.08	1.53	1.91
/i/-/y/	Deviant	-2.64	-2.24	-0.83	1.54	2.62	2.07
Late MMN time window, Before Social Interaction							
Contrast	Condition	LT	RT	FC	CP	LP	RP
/e/-/ε/	Control	-1.62	-1.54	-1.5	0.89	1.66	1.78
/e/-/ε/	Deviant	-1.53	-1.27	-1.38	0.9	1.5	1.94
/i/-/y/	Control	-1.58	-1.46	-1.65	0.4	1.54	2.11
/i/-/y/	Deviant	-1.87	-2.07	-1.82	0.88	2.58	1.95
N100 time window, After Social Interaction							
Contrast	Condition	LT	RT	FC	CP	LP	RP
/e/-/ε/	Control	0.45	-0.27	-0.86	-1.06	0.11	-0.48
/e/-/ε/	Deviant	-0.04	-0.14	-0.78	-0.71	0.05	0.04
/i/-/y/	Control	0.15	0.01	-0.64	-1.34	-0.24	-0.4
/i/-/y/	Deviant	0.24	0.24	-0.69	-1.46	-0.48	-0.44
P200 time window, After Social Interaction							

Contrast	Condition	LT	RT	FC	CP	LP	RP
/e/-/ε/	Control	-0.77	-0.66	0.31	1.08	0.49	0.31
/e/-/ε/	Deviant	-0.95	-0.32	0.51	0.93	0.18	0.33
/i/-/y/	Control	-0.66	-0.45	0.5	0.66	0.32	0.02
/i/-/y/	Deviant	-0.73	-0.24	0.39	0.28	-0.31	-0.09

MMN time window, After Social Interaction							
Contrast	Condition	LT	RT	FC	CP	LP	RP
/e/-/ε/	Control	-2.12	-1.73	-0.95	1.39	1.91	1.73
/e/-/ε/	Deviant	-2.76	-2.07	-1.73	1.51	2.11	2.63
/i/-/y/	Control	-2.01	-2.02	-0.91	1.37	1.77	1.95
/i/-/y/	Deviant	-2.75	-2.34	-0.89	1.48	2.33	2.01

Late MMN time window, After Social Interaction							
Contrast	Condition	LT	RT	FC	CP	LP	RP
/e/-/ε/	Control	-1.56	-1.14	-1.03	1.12	1.47	1.51
/e/-/ε/	Deviant	-1.89	-1.56	-0.99	1.24	1.65	1.73
/i/-/y/	Control	-1.8	-1.33	-1.42	0.75	1.49	2.21
/i/-/y/	Deviant	-2.52	-2.48	-1.76	1.38	2.77	2.3

Table 2. Summary of statistical results on the EEG data over each time window.

N100 time window		
Factors	Before Social Interaction	After Social Interaction
Contrast	$F(1,18) = 4.36, p = .05$	$F(1,18) = 5.81, p = 2.68^{e-2}$
Condition	$F(1,18) = 3.25, p = .09$	$F(1,18) = 1, p = .33$
Site	$F(5,90) = 5.74, p = 9.12^{e-3}$	$F(5,90) = 2.72, p = .09$
Contrast×Condition	$F(1,18) = 0.73, p = .4$	$F(1,18) = 2.17, p = .16$
Contrast×Site	$F(5,90) = 0.5, p = .65$	$F(5,90) = 1.91, p = .15$
Condition×Site	$F(5,90) = 0.54, p = .63$	$F(5,90) = 0.67, p = .53$
Contrast×Condition×Site	$F(5,90) = 1.07, p = .35$	$F(5,90) = 0.64, p = .55$
P200 time window		
Factors	Before Social Interaction	After Social Interaction
Contrast	$F(1,18) = 1.39, p = .25$	$F(1,18) = 5.99, p = 2.49^{e-2}$
Condition	$F(1,18) = 0.62, p = .44$	$F(1,18) = 4.41, p = .05$
Site	$F(5,90) = 2.22, p = .13$	$F(5,90) = 1.91, p = .17$
Contrast×Condition	$F(1,18) = 0.15, p = .7$	$F(1,18) = 6.53, p = 2^{e-2}$
Contrast×Site	$F(5,90) = 0.68, p = .56$	$F(5,90) = 0.76, p = .47$
Condition×Site	$F(5,90) = 1.46, p = .24$	$F(5,90) = 1.5, p = .23$
Contrast×Condition×Site	$F(5,90) = 0.34, p = .71$	$F(5,90) = 0.08, p = .95$
MMN time window		
Factors	Before Social Interaction	After Social Interaction
Contrast	$F(1,18) = 0.21, p = .64$	$F(1,18) = 0.007, p = .99$
Condition	$F(1,18) = 4.93, p = 3.94^{e-2}$	$F(1,18) = 2.5, p = .13$
Site	$F(5,90) = 36.59, p = 7.17^{e-9}$	$F(5,90) = 19.23, p = 4.29^{e-5}$
Contrast×Condition	$F(1,18) = 1.03, p = .32$	$F(1,18) = 0.49, p = .49$
Contrast×Site	$F(5,90) = 1.25, p = .3$	$F(5,90) = 0.49, p = .62$
Condition×Site	$F(5,90) = 1.34, p = .27$	$F(5,90) = 2.42, p = .09$
Contrast×Condition×Site	$F(5,90) = 3.08, p = .05$	$F(5,90) = 1.03, p = .38$
Late time window		
Factors	Before Social Interaction	After Social Interaction
Contrast	$F(1,18) = 1.16, p = .29$	$F(1,18) = 4.15, p = .06$
Condition	$F(1,18) = 2.06, p = .16$	$F(1,18) = 0.44, p = 0.51$
Site	$F(5,90) = 29.71, p = 6.73^{e-9}$	$F(5,90) = 12.87, p = 5.34^{e-4}$
Contrast×Condition	$F(1,18) = 1, p = .76$	$F(1,18) = 0.003, p = .96$
Contrast×Site	$F(5,90) = 1.51, p = .23$	$F(5,90) = 3.49, p = 2.87^{e-2}$
Condition×Site	$F(5,90) = 0.71, p = .48$	$F(5,90) = 5.10, p = 2.73^{e-3}$
Contrast×Condition×Site	$F(5,90) = 2.38, p = .09$	$F(5,90) = 2.2, p = .11$

References

- Atienza, M., & Cantero, J. L. (2001). Complex sound processing during human REM sleep by recovering information from long-term memory as revealed by the mismatch negativity (MMN). *Brain research*, *901*(1-2), 151-160. [https://doi.org/10.1016/s0006-8993\(01\)02340-x](https://doi.org/10.1016/s0006-8993(01)02340-x)
- Armstrong, N., & Pooley, T. (2010). *Social and linguistic change in European French*. Palgrave Macmillan.
- Barrett, K. A., & Fulfs, J. M. (1998). Effect of gender on the mismatch negativity auditory evoked potential. *Journal of the American Academy of Audiology*, *9*(6), 444-451. PMID: 9865777.
- Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., Poeppel, D., & Dikker, S. (2019). Brain-to-Brain Synchrony and Learning Outcomes Vary by Student-Teacher Dynamics: Evidence from a Real-world Classroom Electroencephalography Study. *Journal of Cognitive Neuroscience*, *31*(3), 401-411. https://doi.org/10.1162/jocn_a_01274
- Bhaya-Grossman, I., & Chang, E. F. (2022). Speech computations of the human superior temporal gyrus. *Annual review of psychology*, *73*, 79-102. <https://doi.org/10.1146/annurev-psych-022321-035256>
- Boersma, Paul & Weenink, David (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.20, retrieved 24 September 2022 from <http://www.praat.org/>
- Braun, S., Rosseel, Y., Kempnaers, C., Loas, G., & Linkowski, P. (2015). Self-Report of Empathy: A Shortened French Adaptation of the Interpersonal Reactivity Index (IRI)

- Using Two Large Belgian Samples. *Psychological Reports*, *117*(3), 735-753.
<https://doi.org/10.2466/08.02.PR0.117c23z6>
- Brunellière, A., Dufour, S., Nguyen, N., & Frauenfelder, U.H. (2009). Behavioral and electrophysiological evidence for the impact of regional variation on phoneme perception. *Cognition*, *111*, 390-396. <https://doi.org/10.1016/j.cognition.2009.02.013>
- Brunellière, A., Dufour, S., & Nguyen, N. (2011). Regional differences in the listener's phonemic inventory affect semantic processing: A mismatch negativity (MMN) study. *Brain and language*, *117*(1), 45-51. <https://doi.org/10.1016/j.bandl.2010.12.004>
- Calcus, A., Deltenre, P., Hoonhorst, I., Collet, G., Markessis, E., & Colin, C. (2015). MMN and P300 are both modulated by the featured/featureless nature of deviant stimuli. *Clinical neurophysiology*, *126*(9), 1727-1734. <https://doi.org/10.1016/j.clinph.2014.11.020>
- Chen D, Zhang R, Liu J, Wang P, Bei L, Liu CC, & Li X. (2022). Gamma-band neural coupling during conceptual alignment. *Human Brain Mapping*, *43*(9), 2992-3006.
<https://doi.org/10.1002/hbm.25831>.
- Cheng, H. S., Niziolek, C. A., Buchwald, A., & McAllister, T. (2021). Examining the Relationship Between Speech Perception, Production Distinctness, and Production Variability. *Frontiers in human neuroscience*, *15*, 660948.
<https://doi.org/10.3389/fnhum.2021.660948>
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature neuroscience*, *1*(5), 351-353. <https://doi.org/10.1038/1561>
- Cheour, M., Martynova, O., Näätänen, R., Erkkola, R., Sillanpää, M., Kero, P., Raz, A., Kaipio, M. L., Hiltunen, J., Aaltonen, O., Savela, J., & Hämäläinen, H. (2002). Speech sounds

- learned by sleeping newborns. *Nature*, 415(6872), 599-600.
<https://doi.org/10.1038/415599b>
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition*, 127-149. American Psychological Association. <https://doi.org/10.1037/10096-006>
- Colin, C., Hoonhorst, I., Markessis, E., Radeau, M., de Tourtchaninoff, M., Foucher, A., Collet, G., & Deltenre, P. (2009). Mismatch negativity (MMN) evoked by sound duration contrasts: an unexpected major effect of deviance direction on amplitudes. *Clinical neurophysiology*, 120(1), 51-59. <https://doi.org/10.1016/j.clinph.2008.10.002>
- Conrey, B., Potts, G. F., & Niedzielski, N. A. (2005). Effects of dialect on merger perception: ERP and behavioral correlates. *Brain and Language*, 95(3), 435-449.
<https://doi.org/10.1016/j.bandl.2005.06.008>
- Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review*, 8(2), 144-153. <https://doi.org/10.1177/1754073914558466>
- Davis, M. H. (1980). *Interpersonal Reactivity Index (IRI)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t01093-000>
- Dehaene-Lambertz G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport*, 8(4), 919-924. <https://doi.org/10.1097/00001756-199703030-00021>
- Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: a cross-linguistic study. *Journal of Cognitive Neuroscience*, 12(4), 635-647. <https://doi.org/10.1162/089892900562390>

- Dilley, L.C., & Pitt, M. A. (2007). A study of regressive place assimilation in spontaneous speech and its implications for spoken word recognition. *The Journal of the Acoustical Society of America*, 122(4): 2340-2353. <https://doi.org/10.1121/1.2772226>
- Dufour, S., Nguyen, N., & Frauenfelder, U. H. (2010). Does training on a phonemic contrast absent in the listener's dialect influence word recognition? *The Journal of the Acoustical Society of America*, 128(1), EL43-EL48. <https://doi.org/10.1121/1.3431102>
- Dufour, S., Brunellière, A., & Nguyen, N. (2013). To what extent do we hear phonemic contrasts in a non-native regional variety? Tracking the dynamics of perceptual processing with EEG. *Journal of Psycholinguistic Research*, 42(2), 161-173. <https://doi.org/10.1007/s10936-012-9212-8>
- Dufour, S., Nguyen, N., Pattamadilok, C., & Frauenfelder, U. H. (2016). Does orthographic training on a phonemic contrast absent in the listener's dialect influence word recognition? *The Journal of the Acoustical Society of America*, 140(3), 1871-1877. <https://doi.org/10.1121/1.4962562>
- Eulitz, C., & Lahiri, A. (2004). Neurobiological Evidence for Abstract Phonological Representations in the Mental Lexicon during Speech Recognition. *Journal of Cognitive Neuroscience*, 16(4), 577-583. <https://doi.org/10.1162/089892904323057308>
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *The European journal of neuroscience*, 15(2), 399-402. <https://doi.org/10.1046/j.0953-816x.2001.01874.x>
- Fagyal, Z., Kibbee, D., & Jenkins, F. (2006). *French: A linguistic introduction*. Cambridge University Press.

- Froemke, R. C., Merzenich, M. M., & Schreiner, C. E. (2007). A synaptic memory trace for cortical receptive field plasticity. *Nature*, *450*(7168), 425-429.
<https://doi.org/10.1038/nature06289>
- Gandolfi, G., Pickering, M. J., & Garrod, S. (2023). Mechanisms of alignment: shared control, social cognition and metacognition. *Philosophical transactions of the royal society B*, *378*, 20210362. <http://doi.org/10.1098/rstb.2021.0362>
- Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In *Contexts of accommodation: Developments in applied sociolinguistics* (Vol. 1). Cambridge University Press.
- Goldinger S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, *105*(2), 251-279. <https://doi.org/10.1037/0033-295x.105.2.251>
- Gow, D.W. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics*, *65*, 575-590. <https://doi.org/10.3758/BF03194584>
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95-112. <https://doi.org/10.1007/BF02289823>
- Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, *34*(4), 458-484. <https://doi.org/10.1016/j.wocn.2005.10.001>
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*, *69*(3), 407-22.
<https://doi.org/10.1016/j.neuron.2011.01.019>.
- Kasai, K., Nakagome, K., Iwanami, A., Fukuda, M., Itoh, K., Koshida, I., & Kato, N. (2002). No effect of gender on tonal and phonetic mismatch negativity in normal adults assessed by a

- high-resolution EEG recording. *Cognitive Brain Research*, 13(3), 305-312.
[https://doi.org/10.1016/s0926-6410\(01\)00125-2](https://doi.org/10.1016/s0926-6410(01)00125-2).
- Kazanina, N., Phillips, C., & Idsardi, W. (2006). The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, 103(30), 11381-11386. <https://doi.org/10.1073/pnas.0604821103>
- Kelley, M. C., & Tucker, B. V. (2020). A comparison of four vowel overlap measures. *The Journal of the Acoustical Society of America*, 147(1), 137-145.
<https://doi.org/10.1121/10.0000494>
- Kraus, N., McGee, T., Carrell, T. D., & Sharma, A. (1995). Neurophysiologic bases of speech discrimination. *Ear and hearing*, 16(1), 19-37. <https://doi.org/10.1097/00003446-199502000-00003>
- Lefebvre, A. (1991). *Le français de la région lilloise*. Homme et société (Vol. 18). Éditions de la Sorbonne. <https://doi.org/10.4000/books.pSORbonne.76617>
- Libet, B., Kobayashi, H., & Tanaka, T. (1975). Synaptic coupling into the production and storage of a neuronal memory trace. *Nature*, 258(5531), 155–157.
<https://doi.org/10.1038/258155a0>
- Limesurvey GmbH. (2022). *LimeSurvey: An open source survey tool*.
<https://www.limesurvey.org/>
- Marslen-Wilson, W. D. (1984). Function and process in spoken word recognition: A tutorial review. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 125-150). Hillsdale, NJ: Erlbaum.

- Menning, H., Roberts, L. E., & Pantev, C. (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *Neuroreport*, *11*(4), 817-822.
<https://doi.org/10.1097/00001756-200003200-00032>
- Morton, J. (1979). Word recognition. In J. Morton & J. C. Marshall (Eds.), *Structures and processes* (pp. 108-156). Cambridge, MA: MIT Press.
- Näätänen, R., Paavilainen, P., Tiitinen, H., Jiang, D., & Alho, K. (1993). Attention and mismatch negativity. *Psychophysiology*, *30*(5), 436-450. <https://doi.org/10.1111/j.1469-8986.1993.tb02067.x>
- Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta psychologica*, *42*(4), 313-329.
[https://doi.org/10.1016/0001-6918\(78\)90006-9](https://doi.org/10.1016/0001-6918(78)90006-9)
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*(6615), 432–434. <https://doi.org/10.1038/385432a0>
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical neurophysiology*, *118*(12), 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: a new French lexical database. *Behavior research methods, instruments, & computers*, *36*(3), 516-524.
<https://doi.org/10.3758/bf03195598>
- Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive psychology*, *47*(2), 204-238. [https://doi.org/10.1016/s0010-0285\(03\)00006-9](https://doi.org/10.1016/s0010-0285(03)00006-9)

- Nycz, J., & Hall-Lew, L. (2013). Best practices in measuring vowel merger. In *Proceedings of meetings on acoustics*, 20(1), 060008. <https://doi.org/10.1121/1.4894063>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- Pardo J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119(4), 2382-2393. <https://doi.org/10.1121/1.2178720>
- Pardo J. S. (2013). Measuring phonetic convergence in speech production. *Frontiers in psychology*, 4, 559. <https://doi.org/10.3389/fpsyg.2013.00559>
- Pardo, J. S., Gibbons, R., Suppes, A., Krauss, R. M. (2012). Phonetic convergence in college roommates, *Journal of Phonetics*, 40(1), 190-197, <https://doi.org/10.1016/j.wocn.2011.10.001>.
- Pardo, J. S., Jay, I. C., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, perception & psychophysics*, 72(8), 2254-2264. <https://doi.org/10.3758/bf03196699>
- Partanen, E., Kujala, T., Näätänen, R., Liitola, A., Sambeth, A., & Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. *Proceedings of the National Academy of Sciences of the United States of America*, 110(37), 15145-15150. <https://doi.org/10.1073/pnas.1302159110>
- Peltola, M. S., Kuntola, M., Tamminen, H., Hämäläinen, H., & Aaltonen, O. (2005). Early exposure to non-native language alters pre-attentive vowel discrimination. *Neuroscience letters*, 388(3), 121-125. <https://doi.org/10.1016/j.neulet.2005.06.037>

- Perrin, F., Pernier, J., Bertrand, O., Giard, M. H., & Echallier, J. F. (1987). Mapping of scalp potentials by surface spline interpolation. *Electroencephalography and clinical neurophysiology*, *66*(1), 75-81. [https://doi.org/10.1016/0013-4694\(87\)90141-6](https://doi.org/10.1016/0013-4694(87)90141-6)
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169-190. <https://doi.org/10.1017/S0140525X04000056>
- Pickering, M. J., & Garrod, S. (2021). *Understanding dialogue: Language use and social interaction*. Cambridge University Press. <https://doi.org/10.1017/9781108610728>
- Phillips, C. (2001). Levels of representation in the electrophysiology of speech perception. *Cognitive Science*, *25*(5), 711-731. [https://doi.org/10.1016/S0364-0213\(01\)00049-0](https://doi.org/10.1016/S0364-0213(01)00049-0)
- Quezada, C., Robledo, J. P., Roman, D., & Cornejo, C. (2012). Empathy and pitch convergence. *RLA [online]*. *50* (2), 145-165. <http://dx.doi.org/10.4067/S0718-48832012000200007>.
- Rivera-Gaxiola, M., Csibra, G., Johnson, M. H., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behavioural Brain Research*, *111*(1-2), 13-23. [https://doi.org/10.1016/S0166-4328\(00\)00139-X](https://doi.org/10.1016/S0166-4328(00)00139-X)
- Sato, M. (2020). The neurobiology of sex differences during language processing in healthy adults: A systematic review and a meta-analysis. *Neuropsychologia*, *140*, 107404. <https://doi.org/10.1016/j.neuropsychologia.2020.107404>.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime User 's Guide*. Pittsburgh: Psychology Software Tools, Inc.

- Schweitzer, A., & Lewandowski, N. (2014). Social factors in convergence of F1 and F2 in spontaneous speech. In *Proceedings of the 10th international seminar on speech production* (pp. 391-394).
- Schweitzer, A., & Lewandowski, N. (2013). Convergence of articulation rate in spontaneous speech. In *Proceedings in Interspeech*, 525-529.
<https://doi.org/10.21437/Interspeech.2013-148>
- Shafer, V. L., Schwartz, R. G., & Kurtzberg, D. (2004). Language-specific memory traces of consonants in the brain. *Cognitive Brain Research*, 18(3), 242-254.
<https://doi.org/10.1016/j.cogbrainres.2003.10.007>
- Shestakova, A., Huotilainen, M., Ceponiene, R., & Cheour, M. (2003). Event-related potentials associated with second language learning in children. *Clinical neurophysiology*, 114(8), 1507-1512. [https://doi.org/10.1016/s1388-2457\(03\)00134-2](https://doi.org/10.1016/s1388-2457(03)00134-2)
- Shtyrov, Y., & Pulvermüller, F. (2007). Early MEG activation dynamics in the left temporal and inferior frontal cortex reflect semantic context integration. *Journal of Cognitive Neuroscience*, 19(10), 1633-1642. <https://doi.org/10.1162/jocn.2007.19.10.1633>
- Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences of the United States of America*, 107(32), 14425-14430.
<https://doi.org/10.1073/pnas.1008662107>
- Tamminen, H., Peltola, M. S., Kujala, T., & Näätänen, R. (2015). Phonetic training and non-native speech perception--New memory traces evolve in just three days as indexed by the mismatch negativity (MMN) and behavioural measures. *International journal of psychophysiology*, 97(1), 23-29. <https://doi.org/10.1016/j.ijpsycho.2015.04.020>

- Timm, J., Weise, A., Grimm, S., & Schröger, E. (2011). An Asymmetry in the Automatic Detection of the Presence or Absence of a Frequency Modulation within a Tone: A Mismatch Negativity Study. *Frontiers in psychology, 2*, 189.
<https://doi.org/10.3389/fpsyg.2011.00189>
- Tang, M., Huang, Z. L., Zhong, F., Xiang, J. L., & Wang, X. D. (2020). One-week phonemic training rebuilds the memory traces of merged phonemes in merged speakers. *Brain research, 1740*, 146848. <https://doi.org/10.1016/j.brainres.2020.146848>
- Toufan, R., Aghamolaei, M., & Ashayeri, H. (2021). Differential effects of gender on mismatch negativity to violations of simple and pattern acoustic regularities. *Brain and Behavior, 11*(8), e2248. <https://doi.org/10.1002/brb3.2248>
- Tremblay, K., Kraus, N., Carrell, T. D., & McGee, T. (1997). Central auditory system plasticity: generalization to novel stimuli following listening training. *The Journal of the Acoustical Society of America, 102*(6), 3762-3773. <https://doi.org/10.1121/1.420139>
- Tremblay, K., Kraus, N., & McGee, T. (1998). The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *Neuroreport, 9*(16), 3557-3560. <https://doi.org/10.1097/00001756-199811160-00003>
- Watkins, K., & Paus, T. (2004). Modulation of motor excitability during speech perception: the role of Broca's area. *Journal of Cognitive Neuroscience, 16*(6), 978-987.
<https://doi.org/10.1162/0898929041502616>
- Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csépe, V., Ilmoniemi, R. J., & Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology, 36*(5), 638-642.
<https://doi.org/10.1017/S0048577299981908>

Wu, Y., Guo, X., Gao, Y., Wang, Z., & Wang, X. (2019). Meaning enhances discrimination of merged phonemes: A mismatch negativity study. *Brain research*, 1724, 146433.

<https://doi.org/10.1016/j.brainres.2019.146433>

Yi, H. G., Leonard, M. K., & Chang, E. F. (2019). The Encoding of Speech Sounds in the Superior Temporal Gyrus. *Neuron*, 102(6), 1096-1110.

<https://doi.org/10.1016/j.neuron.2019.04.023>

Yu, D., Ponomarev, A., & Davis, R.L. (2004). Altered representation of the spatial code for odors after olfactory classical conditioning: memory trace formation by synaptic recruitment. *Neuron*, 42, 437-449. [https://doi.org/10.1016/s0896-6273\(04\)00217-x](https://doi.org/10.1016/s0896-6273(04)00217-x)

Zhang, Y., Kuhl, P. K., Imada, T., Iverson, P., Pruitt, J., Stevens, E. B., Kawakatsu, M., Tohkura, Y., & Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: a magnetoencephalography study. *NeuroImage*, 46(1), 226-240.

<https://doi.org/10.1016/j.neuroimage.2009.01.028>

Zhou, W., Wang, Y., & Wang, X (preprint). Short-Term Phonemic Training Helps Phonemic Mergers Form Memory Traces of Merged Phonemes as Revealed by Visual Mismatch Negativity. <http://dx.doi.org/10.2139/ssrn.4496487>.