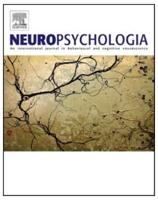
Neuropsychologia 156 (2021) 107831

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: http://www.elsevier.com/locate/neuropsychologia



Neuroplasticity in the phonological system: The PMN and the N400 as markers for the perception of non-native phonemic contrasts by late second language learners

Karin Heidlmayr[a,b,c,d,*], Emmanuel Ferragne[c], Frédéric Isel[d]

[a] *UMR 1253, iBrain, University of Tours, Inserm, Tours, France*
[b] *Max-Planck Institute for Psycholinguistics, Nijmegen, the Netherlands*
[c] *Laboratory CLILLAC-ARP – URP3967, Université de Paris, Paris, France*
[d] *Laboratory Models, Dynamics, Corpus, CNRS/University Paris Nanterre, Paris Lumières, France*

ARTICLE INFO

Keywords:
Second language acquisition
Neuroplasticity
Phonology
Semantic unification
PMN
N400

ABSTRACT

Second language (L2) learners frequently encounter persistent difficulty in perceiving certain non-native sound contrasts, i.e., a phenomenon called “phonological deafness”. However, if extensive L2 experience leads to neuroplastic changes in the phonological system, then the capacity to discriminate non-native phonemic contrasts should progressively improve. Such perceptual changes should be attested by modifications at the neurophysiological level. We designed an EEG experiment in which the listeners’ perceptual capacities to discriminate second language phonemic contrasts influence the processing of lexical-semantic violations. Semantic congruency of critical words in a sentence context was driven by a phonemic contrast that was unique to the L2, English (e.g.,/ɪ/-/iː/, *ship – sheep*). Twenty-eight young adult native speakers of French with intermediate proficiency in English listened to sentences that contained either a semantically congruent or incongruent critical word (e.g., *The anchor of the* ***ship/*sheep*** *was let down*) while EEG was recorded. Three ERP effects were found to relate to increasing L2 proficiency: (1) a left frontal auditory N100 effect, (2) a smaller fronto-central phonological mismatch negativity (PMN) effect and (3) a semantic N400 effect. No effect of proficiency was found on oscillatory markers. The current findings suggest that neuronal plasticity in the human brain allows for the late acquisition of even hard-wired linguistic features such as the discrimination of phonemic contrasts in a second language. This is the first time that behavioral and neurophysiological evidence for the critical role of neural plasticity underlying L2 phonological processing and its interdependence with semantic processing has been provided. Our data strongly support the idea that pieces of information from different levels of linguistic processing (e.g., phonological, semantic) strongly interact and influence each other during online language processing.

1. Introduction

Late second language (L2) learners often experience persistent difficulty in perceiving and producing specific phoneme contrasts in their L2 (e.g., Cutler, 2012). This reduced capacity to discriminate non-native phonemes is often called “phonological deafness” (Polivanov, 1931) (also referred to as “phonological reorganization” (Best et al., 1988; Werker and Tees, 2005, 1984) or “phonological narrowing” (Lewkowicz and Hansen-Tift, 2012)) and illustrated by Troubetzkoy (1939) with the metaphor of “phonological screen”. Interestingly, this selective perception may not be irreversible as suggested by previous phonetic studies reporting evidence of plasticity in the phonological system of adults (Best and Strange, 1992; Flege et al., 1997; Iverson et al., 2012). In particular, sensitivity to non-native sounds can be improved after short-term exposure when listeners directly experience these sounds in a social interaction with native speakers (Kuhl et al., 2003). This indicates that even late learners of a second language are able to learn to perceive and produce phonemes that are not part of their native phoneme repertoire. In the present study, our main goal was to investigate the degree to which the neural basis of the phonological system remains sufficiently malleable beyond early infancy and would thus allow for the acquisition of non-native phonemic contrasts in a second language.

---

[*] Corresponding author. UMR 1253, iBrain, University of Tours, Inserm, Tours, France.
E-mail address: karin.heidlmayr@univ-tours.fr (K. Heidlmayr).

https://doi.org/10.1016/j.neuropsychologia.2021.107831
Received 10 April 2020; Received in revised form 28 December 2020; Accepted 17 March 2021

1.1. Neurocognitive constraints in first and second language acquisition

In language development, during the first year of life, children improve their capacity to discriminate sounds from their native language(s) and begin to lose the capacity to discriminate phonemic contrasts that are not relevant to the language(s) in acquisition (Kuhl, 2004; Sebastian-Galles, 2006; Werker and Tees, 1984). These very early changes in the phonological system are crucial for the later success at language acquisition (Native language neural commitment (NLNC) hypothesis; Kuhl, 2004), but produce the "phonological deafness" effect (Polivanov, 1931), mentioned above. The reasons for this difficulty to discriminate and to learn to discriminate non-native sounds in perception and production (Cutler, 2012; Dupoux et al., 2008; Flege and MacKay, 2004; Sebastián-Gallés et al., 2005) have been argued to reside in a difficulty to create non-native phonological representations later in life, rather than in a perceptual problem (e.g., Dupoux et al., 2008). Nevertheless, the poor discrimination capacity for non-native sound contrasts does not appear to be fixated once and for all. There is evidence for the possibility to improve this capacity with second language experience (Best and Strange, 1992; Flege et al., 1997; Iverson et al., 2012) and even to retune the discrimination capacity of the phonological system from the first to the second language (Ventureyra et al., 2004). In the present study, we aimed to test this plasticity of the phonological system as a function of second language experience.

Late second language acquisition (SLA) has been defined as the acquisition of a language beyond the age of three by some authors (Hakuta, 2003), while others set the critical age much higher, at puberty (Klein, 1996). SLA has been argued, on the one hand, to share some of the general-purpose neural learning mechanisms with first language (L1) acquisition (Hamrick et al., 2018; Klein, 2007). On the other hand, late SLA has also been shown to be determined by considerably different neurocognitive constraints from those involved in early language acquisition (for a review, see Meisel, 2011). Late SLA seems to require more effortful, implicit or explicit, learning strategies in order to integrate the novel language into the neurocognitive language network, which is already committed to the native language (Hernandez and Li, 2007; Morgan-Short et al., 2012). These different preconditions in late SLA are caused by neural maturation, early learning experience and socio-psychological constraints (Kuhl and Rivera-Gaxiola, 2008; Newport et al., 2001). However, the theoretical proposal according to which adult learners of L2 achieve native-like knowledge and proficiency only in rare cases, if at all, has been seriously questioned on the basis of neuroimaging data which reveal that neuroplastic changes underlie second language learning success (see for example, Li and Grant, 2016; lexical level) or by neurophysiological sentence processing data (Caffarra et al., 2015; Morgan-Short, 2014; (morpho)syntactic level; Hahne, 2001; Isel, 2005; syntactic and semantic levels among others). Importantly, over the course of L2 acquisition, quantitative and qualitative shifts in the neurophysiological signature have been observed with developing L2 competence. Higher proficiency and longer L2 exposure as well as earlier L2 age of acquisition and an extensive implicit learning context are often associated with the emergence of more native-like ERP signatures, notably for (morpho)syntactic processing (Caffarra et al., 2015; Morgan-Short, 2014), reflecting an increased automatization of L2 processing (Ullman, 2020). The observation that especially proficiency level and extensive L2 exposure in adult L2 learning can modulate neurophysiological markers in response to syntactic analysis lends support to the theoretical perspective that language use and experience can enable the emergence of grammatical rules and other formal aspects beyond an early age of acquisition (Behrens, 2009; Caffarra et al., 2015; Mitchell et al., 2019). However, the potential changes in the neurophysiological signature of phonological processing and its interaction with lexical-semantic and (morpho)syntactic processes over the course of L2 acquisition are less well known.

1.2. Models of non-native speech perception

What are the neurocognitive mechanisms and changes that would enable late second language learners to handle L2 phonological specificities, at least to some degree? According to the *Perceptual assimilation model* (PAM-L2; Best, 1995; Best and Tyler, 2007) non-native sounds are perceived according to their distance to the articulatory properties of native sounds. The specific L2 sounds are assimilated to already existing phonemic categories created during native language acquisition. In contrast, according to the *Speech learning model* (SLM; Flege, 2002, 1995), novel phonetic categories can be created for non-native phonetic representations, under the condition that a difference between the new L2 phoneme and the closest L1 sound can be perceived by the L2 learner. In other words, the creation of new phonetic categories depends on the cross-language phonetic distance but also on the state of development of L1 phonetic categories in the L2 learner's phonological system, because the L1 phonetic categories become stronger attractors for L2 phonemes the more consolidated they are (Flege and MacKay, 2004). Thus, both PAM-L2 and SLM predict that despite the decreasing capacity to discriminate non-native phonemic contrasts after the first years of life, the phonological system remains plastic to a certain degree, but the two theoretical positions diverge with respect to the mechanisms of this late plasticity (see also, Best and Strange, 1992; Best and Tyler, 2007; Flege et al., 1997).

The originality of the present study was to approach the issue of phonological processing in late second language learners by relying on the mechanism of lexical-semantic integration. More precisely, to examine the extent to which second language learners were sensitive to phonological variation in their L2, we created lexical-semantic incongruities based on phonological changes using a semantic violation paradigm (Kutas and Federmeier, 2011; Kutas and Hillyard, 1980). Concretely, the difference between semantically congruent and incongruent items was implemented by a phonemic contrast that was unique to the L2, English (e.g., /ɪ/-/iː/, *ship – sheep*). Three English phonemic contrasts that are either non phonemic in French – /ɪ/-/iː/ (e.g., ship – sheep) and/s/-/θ/ (e.g., mouse – mouth) – or implemented differently at the phonetic level in the two languages –/r/-/w/ (e.g., rest – west) – were selected. The articulatory-phonetic characteristics of the English /r/ are phonetically quite similar to the French /w/ and can thus lead to a perceptual bias towards the French /w/, which makes the distinction between /r/-/w/ more difficult for French native speakers (Hallé et al., 1999). Concerning the /ɪ/-/iː/ contrast, these vowels are known to be mutually confusable by French learners (Iverson et al., 2012) because there is no phonemic boundary in the French system that would perfectly match the English one. Finally, there is no/s/-/θ/contrast in French, both sounds being often categorized as/s/by French learners (Picard, 2002). According to PAM-L2 (Best, 1995; Best and Tyler, 2007), for these three phonemic contrasts, the phonemes that do not exist in the L1 but only in the L2 would tend to be assimilated to the most similar L1 phoneme, which corresponds to the respectively other phoneme in the current contrasts, by an L2 learner (see also, Cruttenden, 2014, p. 200). If the non-native phoneme is assimilated to the closest L1 phoneme and no distinct phonemic category is established, no clear semantic violation effect should be observed, and this independently of the proficiency level of L2. If, however, with increasing L2 proficiency, new phonemic categories are constructed for non-native phonemes (see also, SLM; Flege, 2002, 1995), the highly proficient L2 learner should be able to detect the semantic violation.

1.3. Experimental evidence for neuroplasticity in the phonological system

Behavioral studies investigating non-native phoneme discrimination have produced heterogeneous results. Broersma and Cutler (2011) observed lexical priming effects that were based on a lack of discrimination between two non-native phonemes, even in proficient L2 speakers. Pallier et al. (1997) showed that native speakers of Spanish who were fluent in Catalan were on average not able to discriminate isolated vowels constituting a phonemic contrast in Catalan. However, there was considerable inter-individual variability in performance and some non-native individuals showed native-like discrimination performance. Indeed, Flege and MacKay (2004) observed that some factors linked to language biography are strong predictors of the success in non-native phonological processing: early L2 AoA and a reduced frequency of L1 usage (higher frequency of L2 usage) were related to a better discrimination of non-native phoneme contrasts. These findings underline the importance of taking into account individual differences in language biography when studying non-native speech processing. Moreover, striking experimental evidence for plasticity in the phonological system during SLA has been provided with specifically targeted phonetic training paradigms that proved successful for improving non-native vowel discrimination at the behavioral level (Iverson et al., 2012).

Moreover, neurophysiological studies have demonstrated an increased neural response in the auditory cortex following a targeted phonological discrimination training. For instance, in a magneto-encephalographical (MEG) study, a training-related increase of the amplitude of the mismatch negativity field (MMF) was observed (Menning et al., 2002). Similarly, in an EEG study, an increase of the amplitudes of an event-related potential (ERP) marker peaking at about 130 ms and localized in the right auditory cortex, and of a second ERP component peaking at around 340 ms and localized in the right anterior superior temporal gyrus and/or inferior prefrontal cortex were found to be related to an increase of the behavioral L2 phoneme discrimination capacity (Alain et al., 2007). Moreover, in an fMRI study on lexical learning of Mandarin Chinese by Dutch native speakers, good learners compared to poor learners showed increased functional connectivity (temporal correlations between the hemodynamic activity of different areas) already before learning onset between areas important for phonological processing, i.e. between the left supplementary motor area and the left precentral gyrus as well as the left insula and the left Rolandic operculum (Veroude et al., 2010). Increased functional connectivity in good compared to poor learners was also found after learning, namely between the left and right supramarginal gyrus. This study demonstrates that pre-existing as well as learning-induced functional connectivity characteristics distinguish good from poor L2 learners (Veroude et al., 2010).

Furthermore, important insight comes from studies investigating structural brain characteristics in bilingualism and their changes with L2 acquisition. L2 proficiency was found to be positively correlated with grey matter (GM) density in the left inferior parietal cortex (Mechelli et al., 2004), whereas age of L2 acquisition (L2 AoA) showed a negative correlation with GM density in this region (Mechelli et al., 2004). Higher L2 AoA was also associated with reinforced white matter connectivity (fractional anisotropy) in the corpus callosum, which is thought to reflect higher processing efficiency with earlier L2 AoA (DeLuca et al., 2019).

Moreover, short-term intense L2 learning was found to be related to an increase in GM density in, among others, the left inferior frontal gyrus (IFG) (Mårtensson et al., 2012; Stein et al., 2012), and to an increase in inter-hemispheric white matter connectivity (Schlegel et al., 2012). Finally, with intensive language learning, L2 proficiency-related laterality shifts in structural connectivity in the perisylvian language network have been observed (Xiang et al., 2015). In less proficient L2 speakers, structural connectivity in the frontal (BA6) to temporal pathway (mainly along the arcuate fasciculus) showed a right hemispheric dominance while with increasing proficiency, a stronger left hemispheric dominance emerged. The stronger right hemispheric dominance in less proficient L2-learners was suggested to reflect the recruitment of additional right-hemisphere areas during phonological processing (Xiang et al., 2015).

To sum up, there is evidence in behavior, neurophysiology as well as in neuroimaging that late second language acquisition and L2 proficiency are associated with functional and structural changes especially in the left-lateralized language network. These neural changes often involve regions that are crucial in phonological processing, among others the superior temporal cortex, and their connectivity with other frontotemporal and parietal regions of the language network, suggesting the critical role of phonological processing and of neural changes in its underlying neural network during L2 learning.

1.4. The present study

In the present electroencephalographic (EEG) study, our main goal was to investigate (1) if the plasticity in the neurocognitive phonological system allows late L2 learners to reach L2 phoneme discrimination capacity, as assessed via neurophysiological language markers (for a review, see Friederici, 2011) that reflect automatic sound deviance detection (phonological mismatch negativity (PMN); Connolly and Phillips, 1994) and phonology-mediated semantic unification (N400; Hagoort et al., 2009; Kutas and Federmeier, 2011) and (2) which factors in the linguistic biography of L2 learners (e.g., frequency of L2 usage, age of L2 acquisition, L2 proficiency, duration of immersion in an L2 environment) may best predict the L2 phonological processing capacity. Here, we combined behavioral and electrophysiological measures in order to obtain insight into the temporally fine-grained online processing of word- and sentence-embedded non-native phonemes. Furthermore, the investigation of the impact of individual differences in the language biography was intended to disentangle critical factors that predict the success in non-native speech processing.

We designed an event-related potential (ERP) experiment in which the capacity of listeners to discriminate second language phonemic contrasts mediated lexical access. This paradigm allowed us to study the perception of non-native phonemic contrasts in a sentential context. The rationale for this methodological choice was to study this question in a more ecological situation of language processing, i.e., with short sentences rather than isolated words or phonemes. In natural language processing, the perceptual system does not deal with isolated phonemic contrasts but on the contrary has to process these contrasts by taking into consideration linguistic information provided by other levels of linguistic analysis. This analysis of the speech signal is realized in a bidirectional way, i.e., in a bottom-up and a top-down manner. A semantic violation paradigm was used in which the difference between semantically congruent and incongruent items was implemented via a phonemic contrast that was unique to the L2, English, (e.g.,/ɪ/-/iː/, *ship – sheep*) but that did not exist in the L1, French. Participants listened to sentences that contained either a semantically congruent item (e.g., *The anchor of the **ship** was let down*) or an incongruent one (e.g., **The anchor of the **sheep** was let down*) and were asked to perform an acceptability judgement.

1.4.1. Phonological mismatch negativity (PMN)

In order to assess if the processing system of L2 learners had developed a native-like capacity to detect deviant phonemes with respect to their predictions based on the sentence context, the phonological mismatch negativity (PMN; also referred to as phonological *mapping* negativity; Connolly and Phillips, 1994; Newman and Connolly, 2009) was considered here as one of the relevant neurophysiological marker of interest. The PMN is a frontal, central to left-lateralized negativity peaking around 200–300 ms after stimulus onset that is sensitive to phonological processing of spoken words/nonwords (Connolly and Phillips, 1994; Kujala et al., 2004; Newman and Connolly, 2009). Interestingly, sometimes the PMN is discussed in terms of an auditory mismatch negativity (MMN). The MMN is a fronto-central negativity peaking at around 150–250 ms after stimulus onset. It has been found to show a task-independent automatic response, in the form of a larger amplitude, in response to deviant (speech) sounds (Garrido et al., 2009; Näätänen et al, 1978, 2007; Pulvermüller and Shtyrov, 2006). The MMN component is typically found in oddball paradigms for individual phonemes or sounds. However, not only perceptual features but also the (linguistic) context and naturalistic occurrence statistics seem to influence this auditory deviance detection. For instance, it has been found to be influenced by information about the probability with which a given sound follows another (transitional probability) (Mittag et al., 2016) and by cognitive and linguistic information, such as semantic or syntactic information (Pulvermüller and Shtyrov, 2006). Furthermore, a late MMN can be elicited by complex auditory stimuli like syllables and words (Czamara et al., 2011). Importantly, the two components, PMN and MMN, differ on several aspects: Whereas the MMN is dependent on the frequency of stimulus occurrence, this is not the case for the PMN (Connolly and Phillips, 1994). Moreover, whereas the neural generators of the PMN have been localized in left fronto-temporal regions (Kujala et al., 2004), the generators of the MMN have been found more posteriorly in the bilateral temporal cortex, involving the primary auditory cortex, and in the prefrontal cortex, left-lateralized in the case of language paradigms (for a review, see Garrido et al., 2009). In the present design, the PMN is a valuable marker to identify if L2 learners process contextualized phonetic contrasts as phonemes belonging to the same or to different phonemic categories.

1.4.2. N400

The second ERP component of interest to – indirectly – assess the learners' capacity to discriminate L2 phonemes in our paradigm was the N400. In ERP research, the N400 component is a negative-going component at posterior sites peaking at around 400 ms after stimulus onset. In psycholinguistics, this ERP component was first shown by Kutas and Hillyard (1980); see also Federmeier and Laszlo, 2009; Kutas and Federmeier, 2011, 2000) to reflect difficulties of lexical semantic integration (e.g., *He spread the warm bread with ****socks***.) during the visual integration of words in English sentences. An N400 effect, i.e., a more negative amplitude in a (semantically) incongruent as compared to a congruent condition is usually found in the time window 200–600 ms after stimulus onset and is largest over centro-parietal sites. The amplitude of the N400 component is thought to be related to semantic properties of words and is sensitive to a range of linguistic and non-linguistic stimuli (Kutas and Federmeier, 2011). Variations in the N400 amplitude have been associated with a variety of neurocognitive functions and processes in different languages, such as lexical-semantic integration in sentences (German: Friederici et al., 1993; French: Isel et al., 2007; English: Kutas and Hillyard, 1980), semantic/conceptual unification (Hagoort, 2017; Hagoort et al., 2009), lexical access (Van Petten and Kutas, 1990), binding (Federmeier and Laszlo, 2009), orthographic/phonological analysis (Deacon et al., 2004), or semantic memory access (Kutas and Federmeier, 2000; Van Berkum, 2009; for reviews, see Kutas and Federmeier, 2011; Lau et al., 2008). The typical semantic N400 is mainly generated by superior and middle temporal, anterior temporal, medial temporal and dorsolateral frontal regions (for a review, see Kutas and Federmeier, 2011). If the processed language stimuli are however presented in the L2 rather than the L1, the N400 may differ from the typical characteristics presented above. For instance, compared to language processing in the L1, the N400 onset and peak latencies can be delayed in the L2 of late second language learners (Isel, 2007; Midgley et al., 2009; Newman et al., 2012; Weber-Fox and Neville, 1996) and the N400 amplitude may be smaller in the L2 (Liang and Chen, 2019; Midgley et al., 2009; Newman et al., 2012; see also, Ito et al., 2017). However, these latency and amplitude differences between L2 and L1 seem to decrease with L2 proficiency (Liang and Chen, 2019; Midgley et al., 2009).

1.4.3. Time-frequency representations (TFR)

Beside ERPs, we were also interested in the oscillatory activity, i.e., the time-frequency representations (TFR) associated with L2 phonological and lexical-semantic processing. In recent years, the investigation of neural oscillations has provided increasing insight into the neural processes underlying prediction and unification in language processing. However, less is known about the oscillatory activity during second language processing. Two frequency bands, namely the beta and lower gamma band, were of particular interest in the present study in order to investigate phonological and semantic processing at the sentence-embedded critical word.

Previously, beta band oscillations (14–30 Hz) have been associated with the maintenance of a status quo in the motor and the cognitive domain (Engel and Fries, 2010; Lewis and Bastiaansen, 2015). During sentence processing, activity in the beta frequency range has been suggested to reflect the active synchronization and maintenance of the current network configuration to represent sentence-level meaning under construction and the propagation of prediction to lower levels of processing (Arnal and Giraud, 2012; Lewis and Bastiaansen, 2015). Semantic violations have previously been found to elicit a beta power decrease in comparison to semantically legal sentences (Kielar et al., 2014; Luo et al., 2010; Wang et al., 2012), in a time window that is very similar to the event-related N400 (Wang et al., 2012). Moreover, meaning processing at the sentence level has also been associated with oscillatory activity in the gamma range, in that low and middle gamma was suggested to reflect the matching of top-down predictions with bottom-up linguistic input (Lewis and Bastiaansen, 2015). In the lower gamma range (30–45 Hz) power reductions were observed for semantic violations compared to legal target words in a sentence context (Penolazzi et al., 2009). Similarly, Wang et al. (2018) observed a correlation between the gamma power during the prediction and the activation periods of predicted target words in a sentence context, which was not the case for unpredicted target words.

With respect to phonological mismatch, previous studies have also associated power modulations in the beta frequency band with the top-down transmission of predictions but also with the rhythmic modulation of sensory sampling (Arnal and Giraud, 2012), with beta power increase being associated with stronger predictions (Scharinger et al., 2016). Moreover, beta power modulations have been found to reflect the strength of categorical speech percepts, in that beta power was positively related to the slope of listeners' psychometric identification function from ambiguous to prototypical phonemes with respect to their phonemic categories (higher beta power was found for the identification of more prototypical phonemes; Bidelman, 2015).

1.4.4. Hypotheses

To our knowledge, this study constitutes the first attempt in the laboratory to investigate the discrimination capacity of L2 phonemes in a more ecological manner, in that it targets the interaction of phonological and lexical-semantic processes in a sentence context. This was done by measuring the neural response (ERP; TFR) in a semantic violation paradigm, in which the semantic violations were (1) *phonologically* mediated vs (2) purely non-phonological *lexical-semantic*, as a control. As for the phonological level of manipulation, if during L2 acquisition new phonemic categories can be established for the non-native phonemes (see also, SLM; Flege, 2002, 1995), L2 learners, especially those who are highly proficient, should be able to detect a phonologically mediated semantic violation. If, however, the non-native phoneme remains assimilated to the closest L1 phoneme and no distinct phonemic category is established (PAM-L2; Best, 1995; Best and Tyler, 2007), no semantic violation effect should be observed (for a discussion on phonetic perceptual discrimination underlying lexical discrimination see, Cutler, 2015). Our main hypotheses were that if late learners can detect the L2 phonemic deviance and the resulting semantic violations, a PMN effect signaling phonological deviance detection and a subsequent N400 effect as a marker of the difficulty of semantic unification should be observed. At the lexical-semantic level of manipulation, a clear N400 effect related to semantic violation processing should be found. Regarding TFR, the detection of a phonological mismatch should be reflected by beta power modulations and the violation of lexical-semantic predictions by beta and low gamma power modulations. Critically, we expected variations of the relevant neurophysiological effects to depend on L2 experience: increased effect sizes at the ERP and TFR levels, as well as better behavioral performance to detect semantic violations were expected for more proficient L2 learners.

2. Methods

2.1. Participants

Twenty-nine right-handed (Edinburgh Handedness Inventory) young adult native speakers of French with intermediate proficiency in English participated in the ERP experiment. One participant was excluded due to technical failure of the stimulus presentation computer during the experiment, resulting in the data of 28 participants (19 female, 9 male; Age: 24.2 ± 3.1 years [18; 30]) being retained for the analysis. All participants were late learners of English who had studied the language at secondary school in France as a foreign language (L2): L2 Age of acquisition: 9.3 ± 2.5 years [3; 14], reported intermediate frequency of daily L2 use: 16.8 ± 12.8% [0.0; 45.0], and showed upper-intermediate proficiency in a standardized test of English as a foreign language (*Certificat de compétences en langues de l'enseignement supérieur* (CLES)): 81.1 ± 15.1% [41.5; 100.0]. Data on linguistic and environmental background measures are displayed in Table 1. By their own account, participants had no history of current or past neurological or psychiatric diseases, they had normal or corrected-to-normal vision and normal color vision as well as no hearing deficiency. Before taking part in the study, the participants were individually informed about the experimental protocol as well as of the procedure of data storage and anonymization. They gave their informed written consent. The data collected were anonymized by applying the European Data FAIR principle (Wilkinson et al., 2016). The study was approved by the local ethics committee of Paris Nanterre University - Department of Psychology and was performed in accordance with the Declaration of Helsinki.

2.2. Stimuli and procedure

A modified version of the semantic violation paradigm (Kutas and Federmeier, 2011; Kutas and Hillyard, 1980) was used in which the difference between semantically congruent and incongruent items was implemented by a phonemic contrast that was unique to the L2, English (e.g.,/ɪ/-/iː/, *ship – sheep*). Three English phonemic contrasts that are not phonemic in French were identified:/r/-/w/(e.g., rest – west),/ɪ/-/iː/(e.g., ship – sheep) and/s/-/θ/(e.g., mouse – mouth). Twenty minimal pairs per contrast were selected and subsequently embedded in a sentential context. Only minimal pairs with both items in the same grammatical class (nouns, verbs, or adjectives) were chosen. For each critical word, i.e., each member of a minimal pair, a congruent sentence context was created, that semantically constrained towards the critical word, constituting the 120 sentences of the phonological - congruent condition (e.g., *The anchor of the **ship** was let down*). For the 120 sentences of the phonological - incongruent condition, the critical word was swapped with the other member of the respective minimal pair (e.g., **The anchor of the **sheep** was let down*). Table 2 displays examples of the different experimental conditions and in the Supplementary Material the complete list of sentences is provided. The underlying idea here was that the lexical-semantic incongruency could only be detected if the listener's perceptive system is sufficiently adapted to the phonological characteristics of English, her/his L2, to discriminate the phonological-lexical manipulation. To test for the lexical (un)predictability of the critical words in each sentence context and consequently to make sure that the congruency manipulation would be perceived as such, a cloze test was first conducted with native speakers of English who did not take part in the further ERP experiment. The cloze probability for each critical word in the respective congruent or incongruent sentence context was assessed in an online questionnaire, in which 14 native speakers of English were requested to complete each sentence fragment preceding the critical word with the first word that came to their mind. The cloze probability was significantly higher for the critical words being embedded in a congruent (cloze prob. 0.39 ± 0.34) than an incongruent (0.00 ± 0.02; $p < .001$) sentence context.

In a control condition, 120 congruent sentences (e.g., *On my way home I stopped by the bakery to get some **bread** for dinner*) and their incongruent counterparts consisting of a purely lexical-semantic violation were created. In the lexical-semantic - incongruent condition, the critical word was replaced by a semantically incongruent word that did not have any phonological similarity (no shared onset phoneme or grapheme, no shared trigram sequence) with the critical word (e.g., **On my way home I stopped by the bakery to get some **gold** for dinner*), but was of the same grammatical class (see Table 2 for examples of the different experimental conditions and the Supplementary Material for the complete list of sentences). The cloze probability at the lexical-semantic level was significantly higher for the critical words embedded in a congruent (0.38 ± 0.34) than an incongruent (0.00 ± 0.00; $p < .001$) sentence context. Critical words at the phonological and lexical-semantic levels were matched for cloze probability (congruent critical words: 0.38 ± 0.34, incongruent critical words: 0.00 ± 0.01, $p < .001$; Linguistic level by Congruency interaction: $F < 1$), spoken frequency (logarithmic) (phonological level: 0.98 ± 0.86; lexical-semantic level: 0.87 ± 0.76; $p > .10$; WEBCELEX (Baayen et al., 1995), COBUILD (Collins Birmingham University International Language Database; Sinclair, 1987)), and word length (number of letters) (phonological level: 4.20 ± 0.75; lexical-semantic level: 4.28 ± 0.80; $F < 1$). All sentences were recorded with natural prosody by a female native speaker of Southern British English in a sound-shielded room. The sound files where then post-processed in Praat 6.0.16 (Boersma and Weenink,

Table 1
Background data. Linguistic and environmental background measures as assessed by a questionnaire are reported. The mean, standard deviation (SD) and range are indicated for each variable.

	Mean	(SD)	Range
Age [years]	24.2	(3.2)	[18; 30]
Frequency of use [%]: L1	79.3	(15.6)	[50.0; 100.0]
Frequency of use [%]: L2	16.8	(13.1)	[0.0; 45.0]
Frequency of use [%]: other languages	3.9	(6.5)	[0.0; 20.0]
AoA L2 [years]	9.3	(2.5)	[3.0; 14.0]
Duration of L2 education [years]	11.7	(2.4)	[7.0; 16.0]
Immersion in L2 environment [months]	3.8	(7.7)	[0.0; 36.0]
L2 proficiency: Grammar [%]	83.3	(17.0)	[36.1; 100.0]
L2 proficiency: Oral comprehension [%]	79.2	(17.3)	[37.5; 100.0]
L2 proficiency: Oral production [%]	70.5	(21.4)	[20.0; 100.0]
L2 proficiency: Total [%]	81.1	(15.4)	[41.5; 100.0]

AoA, Age of Acquisition; L2 proficiency: Sub-tests of a standardized test of English as a foreign language (*Certificat de compétences en langues de l'enseignement supérieur* (CLES)) were used to assess the participants' grammatical competence as well as the oral comprehension and production skills in their L2 English.

Table 2
Examples of the stimuli in the four experimental conditions. The critical word in each sentence is highlighted in bold font. The asterisk indicates the incongruent variant of the critical word. Examples at the phonological level are given in the following order for the three phonemic contrasts:/r/-/w/,/ɪ/-/iː/,/s/-/θ/.

Phonological	Lexical-semantic
Linguistic level	
The bird hurt its **wing**/***ring** and so it couldn't fly away.	He didn't have any cash, so he asked the waiter if payment by **card**/***neck** was possible.
To get a better view of the landscape he walked up the **hill**/***heel** behind the house.	I got soaked in the rain, so I asked him for a towel to **dry**/***skip** my hair.
His lips are swollen because someone punched him in the **face**/***faith** yesterday.	It was so dark in the cave that I had to carry a **lamp**/***cup** in front of my head.

2016): silence segments before sentence onset and after sentence offset were removed and the sound files were intensity-normalized (target intensity: 65 dB). Sounds were presented via DT 770 Pro (Beyerdynamic) headphones. The experimental sessions took place in a dimly lit room. The latency of the critical word onset was identified by auditory analysis and by visual inspection of the spectrogram. Finally, two lists were created, consisting of 50% of the sentences that were from the phonological and 50% from the lexical-semantic level, and half of which were congruent and half incongruent sentences. These lists were presented in a counterbalanced fashion across participants so that each critical word and each sentence was presented only once to each participant and maximally three trials per condition were presented in immediate succession.

The participants listened to sentences that contained either a semantically congruent item (e.g., *The anchor of the **ship** was let down*) or an incongruent one (e.g., **The anchor of the **sheep** was let down*) and were asked to perform an offline acceptability judgement after each stimulus sentence. Each trial started with a fixation cross presented in the center of the screen for 500 ms (see Fig. 1), followed by the auditory presentation of a stimulus sentence, during which the fixation cross continued to be displayed on the screen. Then, solely the fixation cross was presented for 2000 ms, followed by a visual presentation of a question mark, thus prompting the participants to push a button (f/j on the keyboard) to indicate if they considered the sentence acceptable or not. The question mark was displayed until a response was given. Response buttons were counterbalanced across participants but were maintained constant within each participant. Then the fixation cross was displayed again for 2000 ms before the start of the next trial.

2.3. Behavioral data analysis

For the offline acceptability judgement, analyses on accuracy (percent correct responses) and discrimination capacity (d′) were carried out. For discrimination capacity analyses (d′), trials with both correct and incorrect responses were included. The d′ was calculated as $d' = z(\text{Hit rate}) - z(\text{False alarm rate})$, i.e. the difference between the z-transforms of the proportion of incongruent trials to which the participant responded "unacceptable" (Hit rate) and the proportion of congruent trials to which subject responded "unacceptable" (False alarm rate). Statistical analyses were carried out in R version 3.5.1 (R Core Team, 2013): paired *t*-tests were run with function *t.test()* (package *stats v3.5.1*); repeated measures ANOVAs were carried out with function *aov ()* (package *stats v3.5.1*), effect sizes were calculated with function *EtaSq ()* (package *DescTools v0.99.26*), post-hoc tests with functions *lmer()* (package *lme4 v1.1.19*) and *emmeans()* (package *emmeans v1.2.3*); multiple regression analyses were carried out with function *lm()* (package *stats v3.5.1*) and standardized coefficients β were calculated with function *lm.beta()* (package *QuantPsyc v1.5*). For the analysis of the accuracy, a two-way repeated measures ANOVA with Linguistic level (phonological vs. lexical-semantic) and Congruency (congruent vs. incongruent) as within-subjects factors was conducted. For the analysis of the d′, a paired *t*-test with Linguistic level (phonological vs. lexical-semantic) as the within-subjects factor was conducted. Moreover, two multiple regression analyses were conducted on d′ as the dependent variable in order to assess the relative influence of several language background measures on the phonological and lexical-semantic error detection capacity in the L2. Five between-subjects factors were included to analyze their respective part of variance on the d′ at the phonological and the lexical-semantic linguistic levels: age of L2 acquisition (L2 AoA), duration of L2 education, duration of immersion in an L2 environment, frequency of L2 use, and L2 proficiency. A significance level of 0.05 was used for all statistical tests.

2.4. EEG acquisition and preprocessing

EEG data were preprocessed in EEGLAB 13.6.5b (Delorme and Makeig, 2004) and analyzed in FieldTrip (version 20190106) (Oostenveld et al., 2011) within MATLAB (version 9.0.0, R2016a). EEG was recorded from 64 channels that were placed according to the international 10–20 system (Fp1, Fz, F3, F7, FT9, FC5, FC1, C3, T7, TP9, CP5, CP1, Pz, P3, P7, O1, Oz, O2, P4, P8, TP10, CP6, CP2, Cz, C4, T8, FT10, FC6, FC2, F4, F8, Fp2, AF7, AF3, AFz, F1, F5, FT7, FC3, FCz, C1, C5, TP7, CP3, P1, P5, PO7, PO3, POz, PO4, PO8, P6, P2, CPz, CP4, TP8, C6, C2, FC4, FT8, F6, F2, AF4, AF8), mounted in an elastic cap (*ActiCap, Brain Products*), with the vertex electrode (Cz) placed at 50% of the distance between inion and nasion and the left and right ear, and recorded with *Pycorder, Brain Products*. All channels were referenced online against an implicit reference generated by the BrainAmp (*Brain Products*) amplifier. Electrode impedances were kept below 25 kΩ. Data were recorded at a sampling rate of 1000 Hz. An online band-pass filter of 0.01–100 Hz was used. During preprocessing, the data were first filtered offline with a bandpass filter from 0.05 to 45 Hz (Hamming windowed sinc FIR filter) and the continuous data were then cleaned for bursts (using Artifact Subspace Reconstruction - ASR), slow drifts, and strongly movement contaminated time windows, and bad channels were removed (clean_rawdata version 0.31, plugin implemented for EEGLAB) [Maximum tolerated flatline duration: 5 s; Transition band for high-pass filter: [0.25 0.75] Hz; Minimum channel correlation: 0.85; Line Noise Criterion: 4; Standard deviation cutoff for removal of bursts (via ASR): 20; Criterion for removing time windows that were not repaired completely: 0.25]. Next, the continuous data were re-referenced to average reference and then segmented into epochs spanning −1000 ms to +2000 ms relative to critical word onset. Next, an independent component analysis (ICA) decomposition (extended infomax algorithm; Lee et al., 1999) was run, followed by an automatic identification and rejection of artifacted independent components using ADJUST (Mognon et al., 2011), specifically targeting artifacts caused by eye blinks, saccadic eye movements, and muscle activity. Then, by automatic artifact detection on the segmented data, epochs were rejected if amplitudes exceeded ±70 μV (pop_eegthresh, EEGLAB), if linear trends were detected below a maximum slope of 0.5 (pop_rejtrend, EEGLAB), if the joint probability of activity for single as well as all channels exceeded 5 standard deviations (pop_jointprob, EEGLAB), if there was an excessively peaky distribution of activity, that is if the kurtosis of activity for single as well as all channels exceeded 5 standard deviations (pop_rejkurt, EEGLAB), or if there were abnormal spectra in the 0–2 Hz frequency window (maximal deviation from baseline by ±50 dB) to remove residual noise from eye

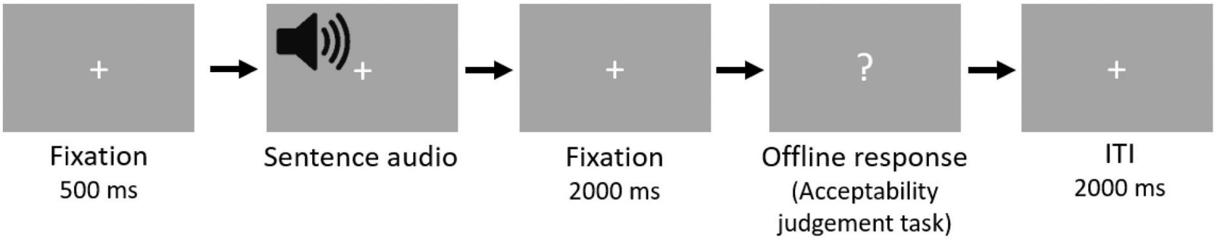


Fig. 1. Timing of a trial.

movement, or in the 20–40 Hz frequency window (maximal deviation from baseline by −100 to +25 dB) to remove residual muscle artifacts (pop_rejspec, EEGLAB). This was followed by the interpolation (spherical) of missing channels. This dataset was subsequently used for time-frequency analysis.

For the event-related brain potential (ERP) analysis, data were further low-pass filtered by 15 Hz (Hamming windowed sinc FIR filter), resegmented to −150 to 1000 ms relative to critical word onset and baseline corrected with the pre-onset time window. After artifact rejection, there were on average 28.5 ± 10.5 (SD) accepted trials per condition when including only trials with correct behavioral responses. On the number of trials remaining for neurophysiological analyses we conducted a two-way repeated measures ANOVA with Linguistic level (phonological vs. lexical-semantic) and Congruency (congruent vs. incongruent) as within-subjects factors. There was a main effect of Linguistic level of analysis ($F(1, 27) = 38.84$, MSE = 23, $p < .001$, $\eta^2_p = .590$), reflecting that a larger number of trials was available for the neurophysiological analyses at the lexical-semantic (31.3 ± 9.4 (SD)) than at the phonological (25.6 ± 10.7 (SD)) level. Moreover, there was a main effect of Congruency ($F(1, 27) = 27.57$, MSE = 49, $p < .001$, $\eta^2_p = .505$), reflecting that a larger number of trials was kept in the congruent (31.9 ± 8.7 (SD)) than in the incongruent (25.0 ± 10.9 (SD)) modality. Finally, there was a Linguistic level of analysis by Congruency interaction ($F(1, 27) = 43.09$, MSE = 16, $p < .001$, $\eta^2_p = .615$), reflecting that the effect of Congruency on the number of kept trials was larger at the phonological (d = 11.9) than at the lexical-semantic (d = 1.9) level. These effects are a trade-off linked to the sensitivity of the design.

2.5. *ERP analysis*

In the current study, critical phonemes that determined the congruency of a word in a given context were not always in word-initial position. This amplifies an already inherent variability in the latency of ERP components. During simple stimulus-locked averaging of trials, part of the relevant signal of a given ERP component might hence get washed out. Alternative approaches have been proposed that take this variability in latency into account, for instance residue iteration decomposition (RIDE; Ouyang et al., 2015a; 2015b). In the present study, RIDE was used to decompose the fully preprocessed and baselined single-trial event-related potential (ERP) data into stimulation- and component-related clusters. Beside the stimulus-related time window spanning 0–50 ms after stimulus onset, two component clusters were included in the analysis. One component was set to an early time window, spanning 50–300 ms after stimulus onset to extract the PMN component, and a second component to a later time window spanning 300–700 ms after stimulus onset to extract the N400 component. After extraction, components were re-baselined with a baseline spanning 0–100 ms after stimulus onset. The selected time windows were based on previous studies (Kutas and Federmeier, 2011; Newman et al., 2012; Weber-Fox and Neville, 1996) but were enlarged to a wider time frame in order to account for temporal variability. At the grand average level, individual-participant component clusters were again subjected to RIDE analysis in order to extract group-level components. In this approach, the component amplitudes are estimated by the covariance between the component template and single trials (or single participants), normalized by the variance of the template. These amplitude estimations were averaged over a subset of six frontal electrodes (Fz, F1, F2, FCz, FC1, FC2) in the case of the early PMN cluster and a subset of six centro-parietal electrodes (Cz, C1, C2, CPz, CP1, CP2) in the case of the later N400 cluster (see also, Fig. 3). These amplitude values were then subjected to statistical analyses.

For the analysis of the amplitude of the PMN and the N400 components, a two-way repeated measures ANOVA with Linguistic level (phonological vs. lexical-semantic) and Congruency (congruent vs. incongruent) as within-subjects factors was conducted. Moreover, multiple regression analyses were conducted on Congruency effect sizes at the phonological and lexical-semantic level, respectively, as the dependent variable in order to assess the relative influence of several language background measures on neurophysiological processes. Five explanatory variables were included to analyze their respective part of variance on the effect size at the phonological and the lexical-semantic levels: age of L2 acquisition (L2 AoA), duration of L2 education, duration of immersion in an L2 environment, frequency of L2 use, and L2 proficiency. A significance level of 0.05 was used for all statistical tests. Statistical analyses were carried out with R version 3.5.1 (R Core Team, 2013): paired *t*-tests were run with function *t.test()* (package *stats v3.5.1*); repeated measures ANOVAs were carried out with function *aov ()* (package *stats v3.5.1*), effect sizes were calculated with function *EtaSq ()* (package *DescTools v0.99.26*), post-hoc tests with functions *lmer()* (package *lme4 v1.1.19*) and *emmeans()* (package *emmeans v1.2.3*); multiple regression analyses were carried out with function *lm()* (package *stats v3.5.1*) and standardized coefficients β were calculated with function *lm_beta()* (package *QuantPsyc v1.5*).

Finally, a cluster-based permutation test using a dependent samples T-statistic was applied to test for an effect of Congruency (congruent vs. incongruent) on the ERP amplitude. This test was carried out separately for each of the two Linguistic levels of analysis (phonological, lexical-semantic) in each of the two time-windows of interest (PMN: 50–300 ms; N400: 300–700 ms). The permutation test involved the Monte Carlo Method to calculate the significance probability, with 1000 draws from the permutation distribution. Cluster correction (maxsum criterion) was used to correct for multiple comparisons. A minimum number of one neighborhood channel needed to show significance in order for a selected sample to be included in the clustering algorithm. An alpha level of 0.05 was chosen for the sample-specific test statistic and an alpha level of 0.05 for the one-sided permutation test (to test only for the negativity effects according to the hypotheses formulated above). Given that one of our main interests lay in the influence of the linguistic background on L2 processing, separate analyses were also carried out for highly and low proficient sub-groups of participants, determined by splitting at the median.

2.6. *Time-frequency analysis*

Statistical analyses on the fully preprocessed electrophysiological data were carried out in FieldTrip (version 20190106) (Oostenveld et al., 2011) in MATLAB (version 9.0.0, R2016a). Time-frequency representations (TFR) were calculated on the single trials between 4 and 44 Hz, in frequency steps of 2 Hz and time steps of 20 ms by using Morlet wavelets with a width of 4 cycles. An alpha level of 0.05 was chosen for the sample-specific test statistic and an alpha level of 0.05 for a one-sided permutation test (corresponding to a threshold of $p < .05$ to detect effects) to detect power decreases in the incongruent compared to the congruent condition. All other aspects of the statistical analyses for TFR effects were identical to the analyses of ERPs. For plotting, TFRs were baseline-corrected by absolute baseline removal spanning the 500 ms preceding critical word onset.

3. Results

3.1. *Behavioral data*

Accuracy. Accuracy data are presented in Fig. 2A. The two-way repeated measures ANOVA including the factors Linguistic level of analysis (phonological vs. lexical-semantic) and Congruency (congruent vs. incongruent) as within-subjects factors and the percent correct responses per condition as the dependent variable revealed a main effect of Linguistic level of analysis ($F(1, 27) = 71.73$, MSE = 69, $p < .001$, $\eta^2_p = .727$), reflecting a higher accuracy at the lexical-semantic (75.63 ± 12.82%) than the phonological (62.29 ± 12.84%) level, as well as a main effect of Congruency ($F(1, 27) = 41.39$, MSE = 232, $p < .001$, $\eta^2_p = .605$), reflecting a higher accuracy in the congruent (78.21 ± 8.99%)

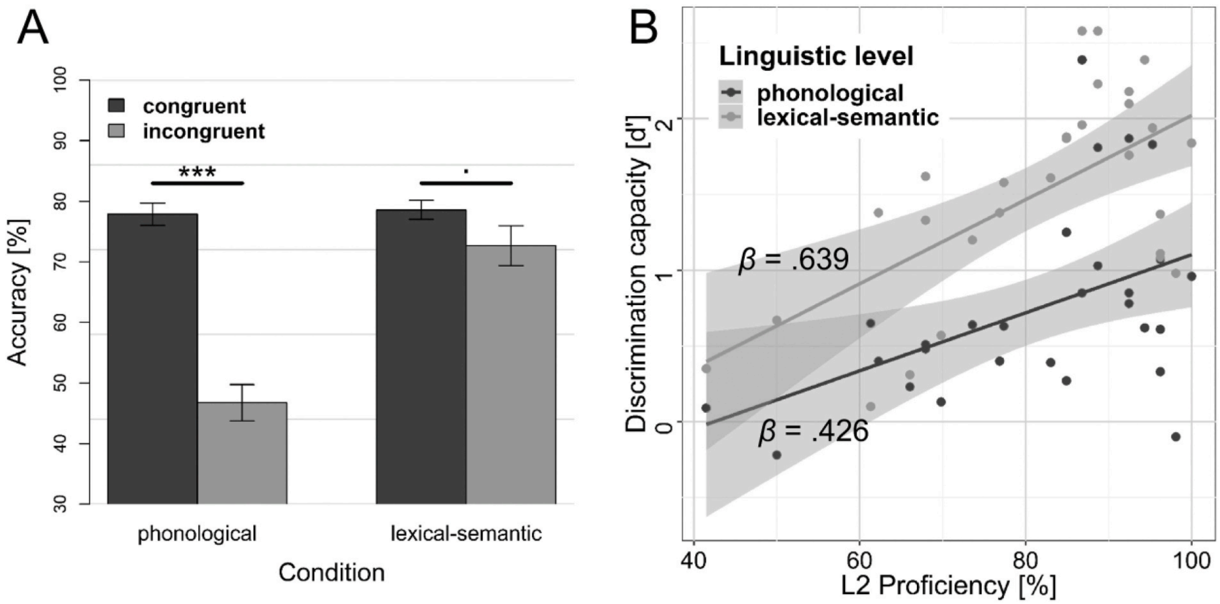


Fig. 2. Behavioral results. A Accuracy of the acceptability judgment in each condition. B Discrimination capacity as calculated via the discrimination capacity measure d′ at both linguistic levels (phonological, lexical-semantic) is plotted as a function of L2 proficiency. ˙ $0.10 > p > .05$, * $p < .05$, ** $p < .01$, *** $p < .001$.

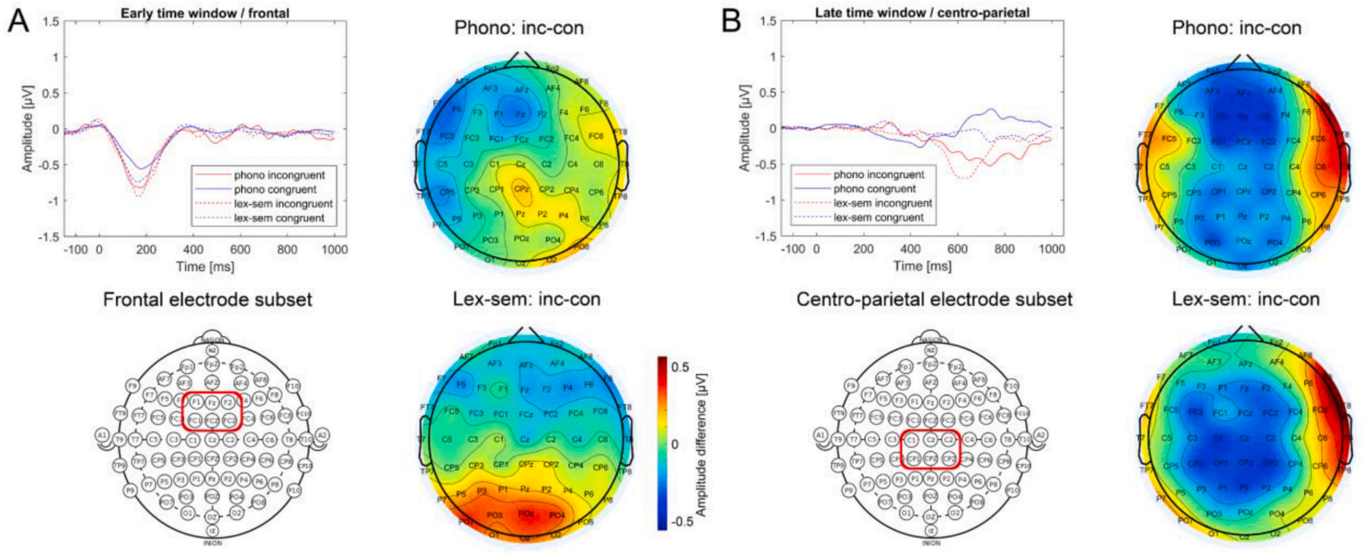


Fig. 3. Event-related potential (ERP) results at the phonological and at the lexical-semantic level. Data were obtained with residue iteration decomposition (RIDE; Ouyang et al., 2015a; 2015b). A Group-averaged ERPs over a subset of six frontal electrodes (Fz, F1, F2, FCz, FC1, FC2) are displayed for the PMN component (left) as well as the topography of amplitude differences for the two contrasts in the time window 150–250 ms (right). The frontal electrode subset used for plotting and for statistical analyses is outlined below. B Group-averaged ERPs over a subset of six centro-parietal electrodes (Cz, C1, C2, CPz, CP1, CP2) are displayed for the N400 component (left) as well as the topography of amplitude differences for the two contrasts in the time window 400–700 ms (right). The centro-parietal electrode subset used for plotting and for statistical analyses is outlined below. inc, incongruent (violation); con, congruent (no violation).

than the incongruent (59.70 ± 16.68%) modality. Moreover, there was a Linguistic level of analysis by Congruency interaction ($F(1, 27) = 71.71$, $MSE = 62$, $p < .001$, $\eta^2_p = .726$), indicating that the effect of Congruency was larger at the phonological ($d = 31.13\%$; $t(40.51) = 9.61$, $p < .001$) than at the lexical-semantic level ($d = 5.89\%$; $t(40.51) = 1.82$, $p > .05$; Fig. 2A).

Discrimination capacity (d′). For the analysis of the discrimination capacity in detecting inacceptable semantic violations, a paired *t*-test with Linguistic level of analysis (phonological vs. lexical-semantic) as a within-subjects factor and d′ as the dependent variable revealed that the discrimination capacity (d′) was better at the lexical-semantic (1.50 ± 0.68) than the phonological level (0.74 ± 0.62; $t(27) = 7.58$, $p < .001$).

A multiple regression analysis was conducted to determine which among the five selected linguistic background variables (age of L2 acquisition (AoA L2), duration of L2 education, duration of immersion in an L2 environment, frequency of L2 use, L2 proficiency) best accounted for the variance of the error discrimination capacity (d′) at the phonological and lexical-semantic linguistic levels in late learners of English. The regression analysis showed that the five predictors explained 46.5% of the variance of the lexical-semantic d′ (Adjusted $R^2 = 0.465$, $F(5,22) = 5.688$, $p < .01$, Cohen's $f^2 = 0.87$; see Table 3). The Cohen's f^2 of 0.87 indicates that this effect is indeed quite large. More specifically, two variables significantly predicted the lexical-semantic d′: L2 proficiency ($\beta = 0.639$, $p < .01$; see Fig. 2B) and duration of L2 education ($\beta = 0.321$,

Table 3

Multiple regression analysis of d′ at the phonological vs. lexical-semantic linguistic levels. Age of L2 acquisition (AoA L2), duration of L2 education, duration of immersion in an L2 environment, frequency of L2 use, and L2 proficiency were included as predictor variables. β, R2 increments (Standardized Coefficient β).

Predictors	phonological d′			lexical-semantic d′		
	β	t	p	β	t	p
Constant		−1.594	.125		−2.737	.012
AoA L2 [years]	.222	1.037	.311	.349	2.064	.051
Duration of L2 education [years]	.200	1.032	.313	.321	2.101	.047
Immersion in L2 environment [months]	.005	.026	.979	-.123	-.794	.436
Frequency of L2 use [%]	.047	.200	.845	.005	.029	.977
L2 proficiency: Total [%]	.426	1.749	.094	.639	3.315	.003

$p < .05$) were positively related to the lexical-semantic d′. In contrast, the model containing these five predictors did overall not significantly explain the variance of the phonological d′ (Adjusted $R^2 = 0.143$, $F(5,22) = 1.902$, $p = .135$).

3.2. ERP data

3.2.1. Early time window: 50–300 ms

RIDE analysis. The two-way repeated measures ANOVA on amplitudes extracted for the early component cluster revealed a significant effect of Congruency reflecting a larger amplitude in the incongruent than the congruent condition over frontal electrodes ($F(1, 27) = 6.42$, MSE $= 0.135$, $p < .05$, $\eta^2_p = .192$; Fig. 3A). No other main effect or interaction was observed (Fs < 1).

Multiple regression analysis. The multiple regression analysis indicated (tendency) that the model of five language background predictors explained 19.0% of the variance of the Congruency effect (PMN effect) at the phonological level (Adjusted $R^2 = 0.190$, $F(5,22) = 2.26$, $p = .083$, Cohen's $f^2 = 0.23$; see Table 4). The Cohen's f^2 of 0.23 indicates that this effect is of medium size. More specifically, one variable significantly predicted the effect size of the PMN: L2 proficiency ($\beta = -0.581$, $p < .05$; see Fig. 4) was inversely related to the phonological effect size, i.e. the higher a participant's proficiency, the smaller the amplitude difference between the incongruent and the congruent condition at the phonological level. In contrast, the model containing these five predictors did overall not significantly explain the variance of the lexical-semantic effect size on this early component (Adjusted $R^2 = -0.054$, $F(5,22) = 0.72$, $p = .612$).

Cluster-based permutation test. Finally, the cluster-based permutation test revealed an effect of Congruency at the phonological level: the incongruent condition showed a larger negativity between 70 and 160 ms over fronto-central, slightly left-lateralized electrodes (Fig. 5A). This effect was driven by the highly proficient subgroup for which a significant effect was found between 55 and 150 ms (Fig. 5B). No other effect was observed in this time window. This early component will in the following be referred to as N100.

Table 4

Multiple regression analysis of the PMN effect size at the phonological level. Age of L2 acquisition (AoA L2), duration of L2 education, duration of immersion in an L2 environment, frequency of L2 use, and L2 proficiency were included as predictor variables. β, R2 increments (Standardized Coefficient β).

Predictors	PMN phono.		
	β	t	p
Constant		1.935	.066
AoA L2 [years]	-.287	−1.380	.182
Duration of L2 education [years]	.232	1.234	.230
Immersion in L2 environment [months]	.230	1.205	.241
Frequency of L2 use [%]	-.040	-.175	.863
L2 proficiency: Total [%]	-.581	−2.452	.023

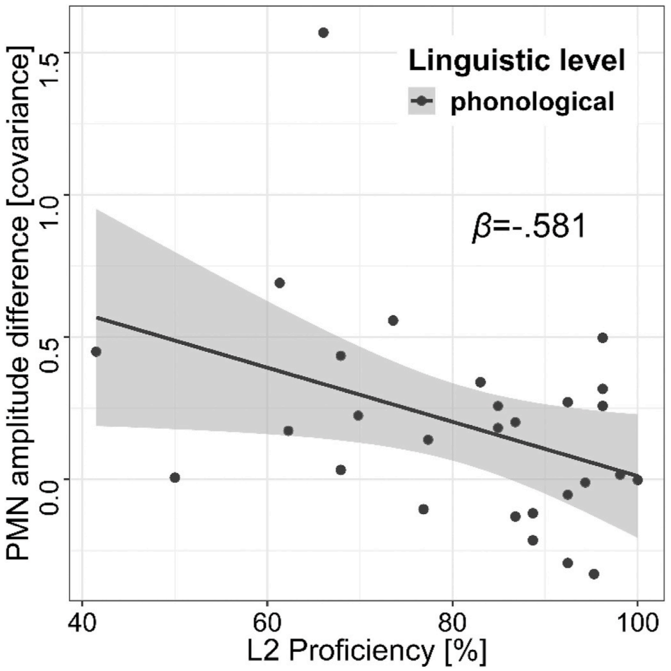


Fig. 4. PMN effect size as a function of L2 proficiency. PMN amplitude effect size at the phonological level is plotted as a function of L2 proficiency. Amplitude was calculated based on covariance with the component template (cf. RIDE; Ouyang et al., 2015a; 2015b).

3.2.2. Late time window: 300–700 ms

RIDE analysis. The two-way repeated measures ANOVA on amplitudes extracted for the later component cluster revealed a significant Congruency by Linguistic level interaction reflecting a larger Congruency effect at the lexical-semantic than the phonological level over central electrodes ($F(1, 27) = 12.93$, MSE $= 0.007$, $p < .001$, $\eta^2_p = .324$; Fig. 3B). No main effect was observed in the ANOVA (ps $> .10$).

Multiple regression analysis. The multiple regression analysis indicated that the model containing the five language background predictors did overall not significantly explain the variance of either the phonological (Adjusted $R^2 = -0.170$, $F(5,22) = 0.21$, $p = .953$) or the lexical-semantic effect size (Adjusted $R^2 = -0.064$, $F(5,22) = 0.68$, $p = .647$) on this later component.

Cluster-based permutation test. Finally, the cluster-based permutation test revealed that for the entire participant group an effect of Congruency (N400 effect) was present at the lexical-semantic level: the incongruent condition was more negative than the congruent one between 520 and 700 ms over central electrodes (Fig. 5C). Moreover, in the highly proficient subgroup an effect of Congruency was present at the phonological level: the incongruent condition was more negative than the congruent one between 575 and 700 ms over centro-parietal electrodes (Fig. 5D). No other effect was observed in this time window.

3.3. Time-frequency data

3.3.1. Early time window: 50–300 ms

No clusters of significant Congruency effects were found.

3.3.2. Late time window: 300–700 ms

A Congruency effect at the lexical-semantic level was found in the beta and lower gamma (20–44 Hz) range in a time window spanning 540–580 ms over central electrodes: beta and low gamma power was reduced in the incongruent compared to the congruent condition at the lexical-semantic level (Fig. 6A and B). Moreover, a second effect was

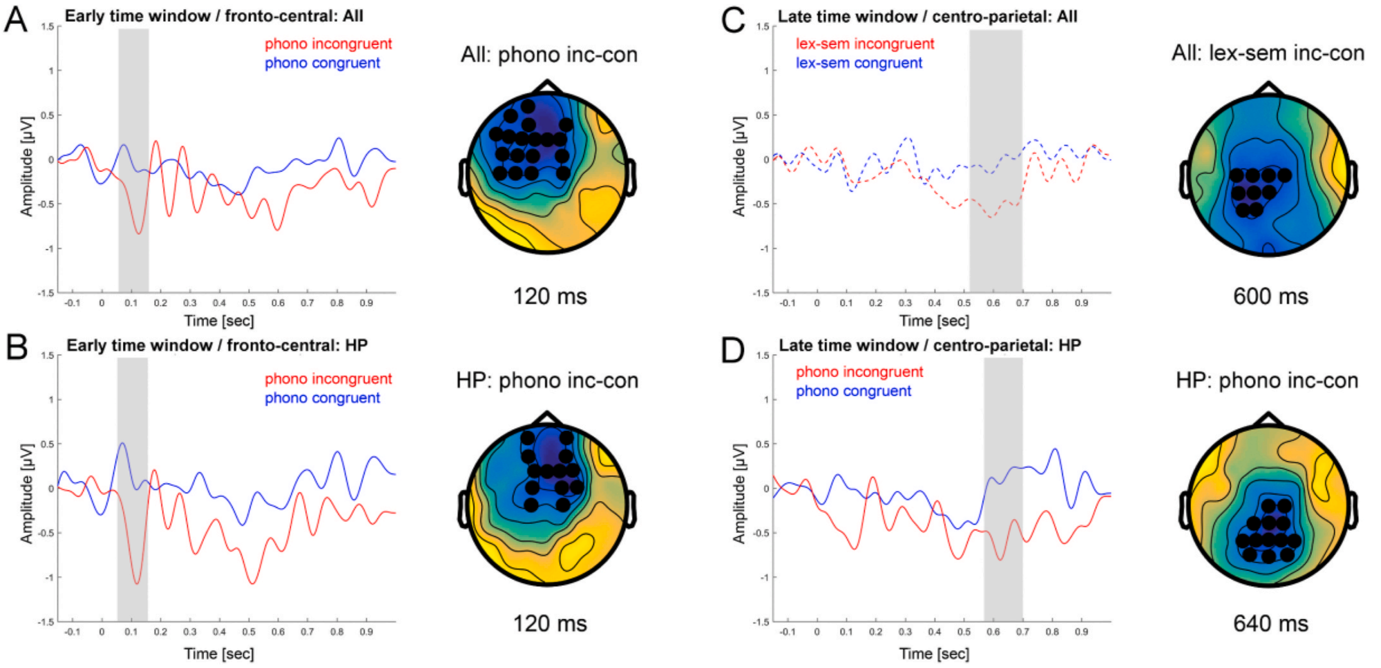


Fig. 5. Event-related potential (ERP) results at the phonological and at the lexical-semantic level. Data were obtained with time-locked cluster-based permutation analyses. A Congruency effect at the phonological level found for all participants in the early time window between 70 and 160 ms over fronto-central electrodes and the topographical plot of the amplitude difference are displayed. B Congruency effect at the phonological level found for highly proficient participants in the early time window between 55 and 150 ms over fronto-central electrodes and the topographical plot of the amplitude difference are displayed. C Congruency effect at the lexical-semantic level found for all participants in the later time window between 520 and 700 ms over centro-parietal electrodes and the topographical plot of the amplitude difference are displayed. D Congruency effect at the phonological level found for highly proficient participants in the later time windows between 575 and 700 ms over centro-parietal electrodes and the topographical plot of the amplitude difference are displayed. All, All participants; HP, Highly proficient participants. Grey bars reflect the time window in which a significant Congruency effect was found. ERP plots for fronto-central visualization are averaged plots over electrodes: Fz, FC1, FC2, Cz, F1, C1, C2, F2. ERP plots for centro-parietal visualization are averaged plots over electrodes: Cz, C1, C2, CPz, CP1, CP2. Topographical distributions represent Congruency effects (difference plots, incongruent minus congruent) at either the phonological or the lexical-semantic level, as specified. Electrodes where effects were observed are marked with a dot.

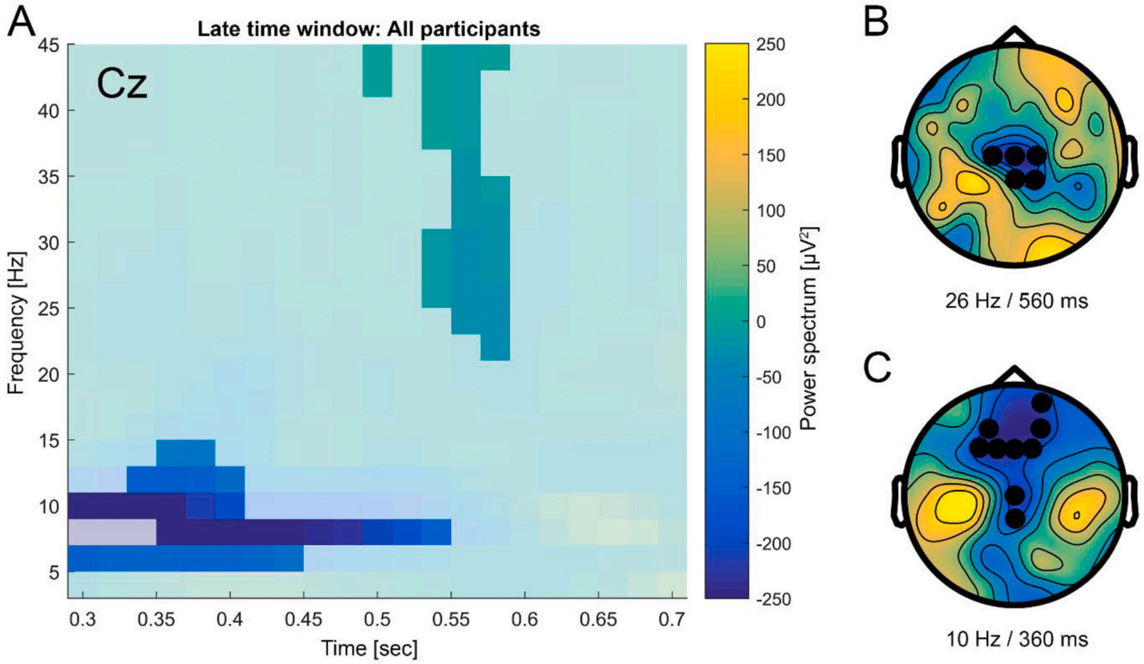


Fig. 6. Time-frequency representation results at the lexical-semantic level. Congruency effects (difference plots, incongruent (violation) minus congruent (no violation)) are presented for the lexical-semantic level. A Congruency effects at the lexical-semantic level for the whole group of participants in the late time window (300–700 ms). B Representative topographical plot for the effect in the beta and gamma band, at 26 Hz and 560 ms after critical word onset. C Representative topographical plot for the effect in the theta and alpha band, at 10 Hz and 360 ms after critical word onset. Electrodes where effects were observed are marked with a dot.

found for this contrast in the higher theta and alpha range (5–15 Hz) in the time window 300–440 ms over fronto-central electrodes: alpha power was reduced in the incongruent compared to the congruent condition at the lexical-semantic level (Fig. 6A and C). No significant effects were found when testing highly and low proficient groups separately.

4. Discussion

In the present study, our main goal was to investigate (1) if the plasticity in the neurocognitive phonological system allows late L2 learners to reach L2 phoneme discrimination capacity, as assessed via neurophysiological language markers that reflect automatic sound deviance detection (phonological mismatch negativity; PMN) and phonology-mediated semantic unification (N400) and (2) which factors in the linguistic biography of L2 learners may best predict the L2 phonological processing capacity. For phonologically mediated semantic violations, we observed an N100, a PMN and an N400 effect, which were however modulated by L2 proficiency: whereas the N100 and the N400 effects were present especially in the highly proficient subgroup, the PMN effect size showed an inverse relation with L2 proficiency. For non-phonological lexical-semantic violations, a more negative deflection in the early time window and a characteristic N400 effect were found. Moreover, the N400 effect at the lexical-semantic level was paralleled by a reduction in alpha power and a slightly later reduction in beta and low gamma power in the incongruent compared to the congruent condition.

4.1. ERP and behavioral data

In the present study, we found behavioral and electroencephalographical evidence that proficiency in late learners of a second language (L2) predicts the capacity to discriminate phonemic contrasts that are specific to the L2. Importantly, the effect size of the phonological mismatch negativity (PMN) found for phonologically mediated semantic violations in a sentence context decreased with L2 proficiency. Moreover, only highly but not low proficient L2 learners showed both an auditory N100 and a characteristic N400 effect in response to phonologically mediated semantic violations in a sentence context. The N400 effect indicates that highly proficient participants processed the semantic violation and therefore that they were able to process the L2-specific phonemic contrast which underlay the semantic violation. In association, the decreasing PMN size with increasing proficiency suggests more automatic, less effortful phonological mapping processes in favor of more profound lexical-semantic processing in the later time window. The higher automaticity of phonological processes seems to be corroborated by the perceptual auditory N100 which was driven by the highly proficient participants. The fact that the PMN pattern was reliably measured with residue iteration decomposition (RIDE; Ouyang et al., 2015a; 2015b) but not with classical time-locked averaging might be due to the particular strength of RIDE to align components with variable timing across trials, which would be mitigated during classical time-locked averaging. Whereas the N100 with very consistent timing and the N400 with a temporally extended waveform may suffer less from classical averaging, this might well have been the case for the PMN. This is especially the case in second language learners who habitually show more variability in their neurophysiological response to L2 stimuli (Isel, 2007; Midgley et al., 2009; Newman et al., 2012; Weber-Fox and Neville, 1996).

The N100 is a negativity peaking around 100 ms (spanning approximately 70–130 ms; Dufour et al., 2013; Sanders et al., 2002) after stimulus onset that is sensitive to physical and temporal characteristics (e.g., change of tonal frequency or intensity, word onset) of the stimulus and dependent on the individual's attentional or arousal state (Näätänen and Picton, 1987; Obleser et al., 2003). The neural generators of the N100 have been localized mainly in the auditory cortex, with additional contributions from the temporal and frontal cortex (Näätänen and Picton, 1987; Näätänen and Winkler, 1999; Obleser et al., 2003). The N100 latency has been found to be more variable in word-embedded non-native phoneme discrimination (Dufour et al., 2013). In a training paradigm using nonwords, Sanders et al. (2002) demonstrated that the N100 amplitude is positively related to the capacity to segment continuous speech into words, independent of acoustic segmentation cues. Moreover, increased selective attention can cause a temporal sharpening of the N100 response (Thornton et al., 2007). In a similar vein, Obleser and Kotz (2011) found an association between N100 amplitudes and comprehension performance in degraded speech, which was suggested to reflect a "bottom-up" processing strategy involving effortful resource allocation. In the present study, the larger auditory N100 in response to the phonemic violation specifically in highly proficient participants may reflect a more native-like, higher temporal precision in L2 auditory processing. Better speech segmentation, as reported by Sanders et al. (2002), may be a core ingredient but the specificity to the phonological violation condition suggests that additional processes are at stake, notably a top-down influence on auditory perception based on linguistic prediction. In the present study, the highly proficient participants may have been able to seize the importance of the phonological information and hence to additionally address more attention to this level of information.

4.1.1. Evidence for new phonemic categories but weaker category boundaries

Previous studies have demonstrated that the sensitivity to non-native phonemic contrasts can increase in late learners of a second language, e.g. with targeted training, but the increase in sensitivity may not reflect the acquisition of a novel phonemic category because the newly acquired phoneme can be integrated in an already existing phonemic category (Dobel et al., 2009). However, the present findings suggest that highly proficient L2 learners perceived the violation induced by a phonemic contrast not as an allophonic variant within a phonemic category but rather as a categorical difference, given that the different lexical-semantic status of the critical word could be identified. Our data corroborate the ideas put forward in the SLM (Flege, 1995, 2002) that the phonological system of late L2 learners still allows for novel phonemic categories to be constructed. However, it remains open whether a novel phonemic category has been created for only one or for both non-native phonemes. Further research should try to identify the role of the relative phonetic proximity of the two non-native phonemes to a native phonemic category (see also, Best and Tyler, 2007; Tyler, 2021) leading to either (1) an assimilation of one of the two L2 phonemes, i.e., the phonetically closer one, to the closest L1 phoneme and the creation of a new category for only the phonetically more distant phoneme, or (2) the creation of new phonemic categories for both non-native phonemes.

Interestingly, the L2 speakers showed good behavioral performance in accepting correct sentences (i.e., no phonologically mediated semantic violation) but their performance was weaker in rejecting unacceptable sentences. These findings may indicate a gradual progression in the creation of non-native phonemic categories. That is, whereas the contrast between the L2 phoneme prototypes seems to be established already at low to intermediate L2 proficiency, the sharpening of the fine-grained categorization boundary which is relevant for categorizing more variable realizations of the phoneme (e.g., word-embedded realizations of the phoneme) may need more time and experience during L2 acquisition. The high proportion of accurate responses in the correct condition and the weaker performance in the violation condition supports this interpretation, potentially reflecting a good handling when a concrete phonemic realization closely matches the expected prototype but greater difficulty to reject less prototypical realizations. Hence, the difficulty might lie rather in the sharpness and delimitation of the phonemic category boundaries than in the capacity to establish a new phonemic category. This weaker boundary delimitation may also underlie previously observed priming effects across phonemic categories even in highly proficient L2 speakers, but not in native speakers (Broersma and Cutler, 2011). Future research should investigate the sharpening of category boundaries with increasing L2 proficiency and experience, as well as changes in the underlying neural processing dynamics.

4.2. Time-frequency data

Power modulations in the beta-gamma (22–44 Hz) and theta-alpha (5–15 Hz) range in response to purely lexical-semantic violations were observed, independent of proficiency. Beta band oscillations (14–30 Hz) have previously been associated with the active synchronization and maintenance of the current neurocognitive state to represent sentence-level meaning under construction and the propagation of prediction to lower levels of processing (Arnal and Giraud, 2012; Lewis and Bastiaansen, 2015). Moreover, meaning-processing at the sentence level has been associated with oscillatory activity in the gamma range (30–45 Hz), in that low and middle gamma was suggested to reflect the matching of top-down predictions with bottom-up linguistic input (Lewis and Bastiaansen, 2015). The present finding of a power decrease in the beta and lower gamma range for semantically incongruent compared to congruent words (with respect to the predictions based on the preceding sentence context) is in accordance with previous findings and corroborates the idea that beta oscillations play a role during the construction of sentence meaning. The beta power decrease for semantic violations may hence reflect a disruption and reconfiguration of this process. Similarly, the idea that lower gamma oscillations play a role in the matching between semantic predictions and the bottom-up input is corroborated by the current data, with gamma power decreases indicating a mismatch between predictions and input. Finally, the power decrease observed in the theta-alpha range (5–15 Hz) for semantically unexpected compared to congruent words is also consistent with previous findings (Kielar et al., 2014) and may reflect semantic unification involving additional domain general, top-down cognitive control processes (see also, Klimesch, 2012; Lam et al., 2016; Rihs et al., 2007; Terporten et al., 2019). Overall, in late L2 learners, the capacity to identify lexical-semantic violations in a sentence context is reflected by characteristic, native-like oscillatory patterns that are associated with sentence-level semantic unification processes, under inclusion of additional attentional and control processes.

With respect to phonological mismatch processing, previous studies have also associated power modulations in the beta frequency band with the top-down transmission of predictions at the phonological level (Arnal and Giraud, 2012), with beta power increases being associated with stronger phonological predictions (Scharinger et al., 2016). In the present study, no beta power modulations for phonemic violations were found in either the highly or low proficient group. Although our expectations were based on previous findings in the native language processing literature, similar dynamics were expected for non-native speech processing at least in highly proficient participants. Research comparing oscillatory dynamics in native and non-native speech is relatively recent but there is evidence that in non-native speech processing, oscillatory dynamics are more variable than in native speech processing (Drijvers et al., 2019; Jin et al., 2014; Lewis et al., 2016; Pérez et al., 2015). This may be due to a less automatized and slower integration of semantic and phonological information as well as a higher cognitive load (Pérez et al., 2015), the need for additional attentional focus (Lewis et al., 2016), or the involvement of different processing strategies (Drijvers et al., 2019) in non-native compared to native speech and information processing.

Here, the cognitive load linked to a reduced automaticity in integrating phonological and lexical-semantic information in the L2 might have attenuated the oscillatory signal for this manipulation, even in the highly proficient group. In an alternative study design, providing instructions or an attentional cue that would allow to focus specifically on the critical part of speech could potentially have a facilitatory role and allow for more profound phonological and semantic processing in the phonologically mediated lexical-semantic manipulation. However, paralleling the robust ERP N400 effect at the lexical-semantic level, a TFR effect was found, indicating that native-like oscillatory markers can emerge for L2 processing. Further research, potentially involving a longitudinal study design, should shed further light on the quantitative and qualitative changes of TFR signatures during L2 acquisition and investigate how these changes parallel the evolution of ERP signatures.

4.3. The role of individual differences in the language biography

Concerning the language background measures (age of L2 acquisition (AoA L2), duration of L2 education, duration of immersion in an L2 environment, frequency of L2 use, L2 proficiency), some factors were more reliable predictors for perceptual sensitivity at the behavioral level than others. That is, especially L2 proficiency – and to a minor degree the duration of L2 education - seems to be a strong predictor for the accuracy of L2 comprehension, as assessed here via the accuracy to discriminate phonological and lexical-semantic violations. Interestingly, it has previously been found that with intense training to discriminate L2 phonemic contrasts, sensory and perceptual plasticity emerge at different timescales. For example, whereas perceptual identification for Mandarin lexical tones was already found earlier during learning, evidence for sensory encoding in the form of neural tracking of the F0 tone contour (frequency-following response; FFR) was only found at a very advanced level of behavioral L2 phonemic discrimination performance (Reetzke et al., 2018). In the same vein, it was found that individuals with better performance in L2 phonological perception showed reduced inter-trial FFR variability (Omote et al., 2017). In the present study, higher L2 proficiency was found to be associated with better behavioral discrimination scores and, at the neurophysiological level, with a larger auditory N100 effect (probably reflecting more fine-tuned timing of the neural response), a smaller PMN effect and a larger effect on the subsequent semantic N400. These findings suggest profound neural changes during second language learning not only in the substrate underlying sensory, phonological or semantic levels but importantly also in the connection between these levels of processing in the language network.

Which are the parameters that enable an L2 learner to reach high proficiency at the different levels of linguistic processing and more specifically at the phonological level? At the phonological level, it has been argued that implicit learning may not be sufficient to establish L2 phonemic discrimination capacity (Dobel et al., 2009). In contrast, there is evidence that intense phonological production training (guided learning) may lead to increased performance in both L2 production and comprehension at the phonemic level (Iverson et al., 2012; Macedonia, 2013), up to the prosodic level (Guyot-Talbot et al., 2016; see also, *motor theory of speech;* Galantucci et al., 2006; Liberman et al., 1967). Furthermore, it has been argued that late L2 learning, and especially phonological learning, more strongly relies on top-down processes, involving attention, relative to bottom-up statistical and implicit learning (White et al., 2013). However, the phonological learning success appears to also strongly depend on the type of phonemic contrast involved (Best et al., 1988; Dobel et al., 2009; Dupoux et al., 2008; Golestani and Zatorre, 2009). In general, it seems to be beneficial to direct an attentional focus on particularly difficult elements of the new language to acquire, be it a syntactic construction or a phonemic contrast that is specific to the L2 and not yet known to the learner from the L1 (e.g., Dobel et al., 2009). Thus, it may be helpful to work towards learning and teaching methods that explicitly involve L2 production as well as explicit focalization on the phonological features to be acquired in order to improve phonological processing during SLA.

Finally, the capacity to acquire non-native phonology has previously been shown to also be of considerable inter-individual variability related to certain neuroanatomical features (Golestani et al., 2006; Golestani and Zatorre, 2009; Omote et al., 2017; Pruitt et al., 2006). Thus, future research should try to identify the facilitating and limiting factors for neuroplastic changes in the phonological system that reside in neurocognitive individual differences as well as in environmental factors such as the language biography and the conditions of second language acquisition.

4.4. Conclusion

The present findings corroborate the idea that neuronal plasticity in the human brain allows for second language acquisition beyond native language attunement early in life. This seems to be the case even for seemingly hard-wired linguistic features such as the discrimination of phonemic contrasts. In the current study, we were particularly interested in the interaction between phonological and semantic levels of processing. To our knowledge, this is the first time that behavioral and neurophysiological evidence for the neural plasticity underlying these linguistic processes and especially for their strong interdependence has been provided. However, further research is necessary to identify the facilitating and limiting factors for neuroplastic changes in the phonological system that reside in neurocognitive individual differences as well as in environmental factors such as the language biography and the conditions of second language acquisition. Finally, our data strongly support the idea that pieces of information from different levels of linguistic processing (e.g., phonological, semantic) heavily interact and influence each other during online language processing.

Credit author statement

Karin Heidlmayr: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; Emmanuel Ferragne: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing – review & editing; Frédéric Isel: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

None.

Acknowledgments

We are very grateful to Hannah King for her support in stimulus creation, to Irena Ilic and Mahsa Bahrami for data collection, to Maud Pélissier for the management of the EEG equipment and testing facilities and to the Labex “Empirical Foundations of Linguistics” (EFL) [ANR-10-LABX-0083] for the provision of the EEG recording system. This work was supported by the project “SOPHOCLE” - IDEX Sorbonne Paris Cité University and by the ANR program “Investissements d’Avenir” - Labex EFL [ANR-10-LABX-0083].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2021.107831.

References

Alain, C., Snyder, J.S., He, Y., Reinke, K.S., 2007. Changes in auditory cortex parallel rapid perceptual learning. Cerebr. Cortex 17, 1074–1084. https://doi.org/10.1093/cercor/bhl018.

Arnal, L.H., Giraud, A.L., 2012. Cortical oscillations and sensory predictions. Trends Cognit. Sci. 16, 390–398. https://doi.org/10.1016/j.tics.2012.05.003.

Baayen, H., Piepenbrock, R., Gulikers, L., 1995. The CELEX Lexical Database [webcelex]. Linguistic Data Consortium. University of Pennsylvania, Philadelphia, PA, USA.

Behrens, H., 2009. Usage-based and emergentist approaches to language acquisition. Linguistics 47. https://doi.org/10.1515/LING.2009.014.

Best, C.T., 1995. A direct realist view of cross-language speech perception. In: Strange, W. (Ed.), Speech Perception and Linguistic Experience: Issues in Cross-Language Research. York Press, Timonium, MD, pp. 171–204.

Best, C.T., McRoberts, G.W., Sithole, N.M., 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. J. Exp. Psychol. Hum. Percept. Perform. 14, 345–360. https://doi.org/10.1037//0096-1523.14.3.345.

Best, C.T., Strange, W., 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. J. Phonetics 20, 305–330.

Best, C.T., Tyler, M., 2007. Nonnative and second-language speech perception: commonalities and complementarities. In: Bohn, O.S., Munro, M. (Eds.), Second-Language Speech Learning: the Role of Language Experience in Speech Perception and Production. A Festschrift in Honour of James E. Flege. Amsterdam, pp. 13–34.

Bidelman, G.M., 2015. Induced neural beta oscillations predict categorical speech perception abilities. Brain Lang. 141, 62–69. https://doi.org/10.1016/j.bandl.2014.11.003.

Boersma, P., Weenink, D., 2016. Praat: a System for Doing Phonetics by Computer, version 6.0. 16.

Broersma, M., Cutler, A., 2011. Competition dynamics of second-language listening. Q. J. Exp. Psychol. 64, 74–95. https://doi.org/10.1080/17470218.2010.499174.

Caffarra, S., Molinaro, N., Davidson, D., Carreiras, M., 2015. Second language syntactic processing revealed through event-related potentials: an empirical review. Neurosci. Biobehav. Rev. 51, 31–47. https://doi.org/10.1016/j.neubiorev.2015.01.010.

Connolly, J.F., Phillips, N.A., 1994. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. J. Cognit. Neurosci. 6, 256–266. https://doi.org/10.1162/jocn.1994.6.3.256.

Cruttenden, A., 2014. Gimson’s Pronunciation of English. Taylor & Francis.

Cutler, A., 2015. Representation of second language phonology. Appl. Psycholinguist. 36, 115–128. https://doi.org/10.1017/S0142716414000459.

Cutler, A., 2012. Native Listening: Language Experience and the Recognition of Spoken Words. The MIT Press, Cambridge, MA.

Czamara, D., Bruder, J., Becker, J., Bartling, J., Hoffmann, P., Ludwig, K.U., Müller-Myhsok, B., Schulte-Körne, G., 2011. Association of a rare variant with mismatch negativity in a region between KIAA0319 and DCDC2 in dyslexia. Behav. Genet. 41, 110–119. https://doi.org/10.1007/s10519-010-9413-6.

Deacon, D., Dynowska, A., Ritter, W., Grose-Fifer, J., 2004. Repetition and semantic priming of nonwords: implications for theories of N400 and word recognition. Psychophysiology 41, 60–74. https://doi.org/10.1111/1469-8986.00120.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.

DeLuca, V., Rothman, J., Bialystok, E., Pliatsikas, C., 2019. Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. Proc. Natl. Acad. Sci. Unit. States Am. 116, 7565–7574. https://doi.org/10.1073/pnas.1811513116.

Dobel, C., Lagemann, L., Zwitserlood, P., 2009. Non-native phonemes in adult word learning: evidence from the N400m. Phil. Trans. Biol. Sci. 364, 3697–3709. https://doi.org/10.1098/rstb.2009.0158.

Drijvers, L., van der Plas, M., Özyürek, A., Jensen, O., 2019. Native and non-native listeners show similar yet distinct oscillatory dynamics when using gestures to access speech in noise. Neuroimage 194, 55–67. https://doi.org/10.1016/j.neuroimage.2019.03.032.

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. J. Psycholinguist. Res. 42, 161–173. https://doi.org/10.1007/s10936-012-9212-8.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., Peperkamp, S., 2008. Persistent stress ‘deafness’: the case of French learners of Spanish. Cognition 106, 682–706. https://doi.org/10.1016/j.cognition.2007.04.001.

Engel, A.K., Fries, P., 2010. Beta-band oscillations—signalling the status quo? Curr. Opin. Neurobiol. 20, 156–165. https://doi.org/10.1016/j.conb.2010.02.015.

Federmeier, K.D., Laszlo, S., 2009. Time for meaning: electrophysiology provides insights into the dynamics of representation and processing in semantic memory. In: Ross, B. H. (Ed.), The Psychology of Learning and Motivation. Elsevier, San Diego, CA, pp. 1–44.

Flege, J.E., 2002. Interactions between the native and second-language phonetic systems. In: Burmeister, P., Piske, T., Rohde, A. (Eds.), An Integrated View of Language Development: Papers in Honor of Henning Wode. Trier, Germany, pp. 217–243.

Flege, J.E., 1995. Second-language speech learning: findings, and problems. In: Strange, W. (Ed.), Speech Perception and Linguistic Experience: Theoretical and Methodological Issues. York Press, Timonium, MD.

Flege, J.E., Bohn, O.-S., Jang, S., 1997. Effects of experience on non-native speakers’ production and perception of English vowels. J. Phonetics 25, 437–470. https://doi.org/10.1006/jpho.1997.0052.

Flege, J.E., MacKay, I.R., 2004. Perceiving vowels in a second language. Stud. Sec. Lang. Acquis. 26, 1–34.

Friederici, A.D., 2011. The brain basis of language processing: from structure to function. Physiol. Rev. 91, 1357–1392. https://doi.org/10.1152/physrev.00006.2011.

Friederici, A.D., Pfeifer, E., Hahne, A., 1993. Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. Cognit. Brain Res. 1, 183–192. https://doi.org/10.1016/0926-6410(93)90026-2.

Galantucci, B., Fowler, C.A., Turvey, M.T., 2006. The motor theory of speech perception reviewed. Psychonomic Bull. Rev. 13, 361–377. https://doi.org/10.3758/BF03193857.

Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009. The mismatch negativity: a review of underlying mechanisms. Clin. Neurophysiol. 120, 453–463. https://doi.org/10.1016/j.clinph.2008.11.029.

Golestani, N., Molko, N., Dehaene, S., LeBihan, D., Pallier, C., 2006. Brain structure predicts the learning of foreign speech sounds. Cerebr. Cortex 17, 575–582. https://doi.org/10.1093/cercor/bhk001.

Golestani, N., Zatorre, R.J., 2009. Individual differences in the acquisition of second language phonology. Brain Lang. 109, 55–67. https://doi.org/10.1016/j.bandl.2008.01.005.

Guyot-Talbot, A., Heidlmayr, K., Ferragne, E., 2016. Entraînements à la prosodie des questions ouvertes et fermées de l'anglais chez des apprenants francophones. In: Presented at the Journée d'études sur la parole, Paris, France. https://hal-univ-diderot.archives-ouvertes.fr/hal-01371892.

Hagoort, P., 2017. The core and beyond in the language-ready brain. Neurosci. Biobehav. Rev. https://doi.org/10.1016/j.neubiorev.2017.01.048.

Hagoort, P., Baggio, G., Willems, R.M., 2009. Semantic unification. In: Gazzaniga, M.S. (Ed.), The Cognitive Neurosciences. MIT Press, Boston, pp. 819–836.

Hahne, A., 2001. What's different in second-language processing? Evidence from event-related brain potentials. J. Psycholinguist. Res. 30, 251–266.

Hakuta, K., 2003. Bilingualism and multilingualism. International Encyclopedia of Linguistics.

Hallé, P., Best, C.T., Levitt, A.G., 1999. Phonetic vs. phonological influences on French listeners: perception of American English approximants. J. Phonetics 27, 281–306.

Hamrick, P., Lum, J.A.G., Ullman, M.T., 2018. Child first language and adult second language are both tied to general-purpose learning systems. Proc. Natl. Acad. Sci. Unit. States Am. 115, 1487–1492. https://doi.org/10.1073/pnas.1713975115.

Hernandez, A.E., Li, P., 2007. Age of acquisition: its neural and computational mechanisms. Psychol. Bull. 133, 638–650. https://doi.org/10.1037/0033-2909.133.4.638.

Isel, F., 2007. Syntactic and referential processes in second-language learners: event-related brain potential evidence. Neuroreport 18, 1885–1889. https://doi.org/10.1097/WNR.0b013e3282f1d518.

Isel, F., 2005. Special issue in honour of Elizabeth Bates. First- and second-language processing: cross-linguistic neurophysiological evidence. In: Editions Modulaires Européennes, Langage et l'homme. Cortil-Wodon, Belgium, pp. 79–95.

Isel, F., Hahne, A., Maess, B., Friederici, A.D., 2007. Neurodynamics of sentence interpretation: ERP evidence from French. Biol. Psychol. 74, 337–346. https://doi.org/10.1016/j.biopsycho.2006.09.003.

Ito, A., Martin, A.E., Nieuwland, M.S., 2017. On predicting form and meaning in a second language. J. Exp. Psychol. Learn. Mem. Cognit. 43, 635–652. https://doi.org/10.1037/xlm0000315.

Iverson, P., Pinet, M., Evans, B.G., 2012. Auditory training for experienced and inexperienced second-language learners: native French speakers learning English vowels. Appl. Psycholinguist. 33, 145–160. https://doi.org/10.1017/S0142716411000300.

Jin, Y., Diaz, B., Colomer, M., Sebastian-Galles, N., 2014. Oscillation encoding of individual differences in speech perception. PloS One 9, e100901. https://doi.org/10.1371/journal.pone.0100901.

Kielar, A., Meltzer, J.A., Moreno, S., Alain, C., Bialystok, E., 2014. Oscillatory responses to semantic and syntactic violations. J. Cognit. Neurosci. 1–23. https://doi.org/10.1162/jocn_a_00670.

Klein, W., 2007. Mechanismen des Erst- und Zweitspracherwerbs. Sprache - Stimme - Gehör 31, 138–143. https://doi.org/10.1055/s-2007-985818.

Klein, W., 1996. Language acquisition at different ages. In: Magnusson, D. (Ed.), Individual Development over the Lifespan: Biological and Psychosocial Perspectives. Cambridge University Press, Cambridge, pp. 88–108.

Klimesch, W., 2012. Alpha-band oscillations, attention, and controlled access to stored information. Trends Cognit. Sci. 16, 606–617. https://doi.org/10.1016/j.tics.2012.10.007.

Kuhl, P., Rivera-Gaxiola, M., 2008. Neural substrates of language acquisition. Annu. Rev. Neurosci. 31, 511–534. https://doi.org/10.1146/annurev.neuro.30.051606.094321.

Kuhl, P.K., 2004. Early language acquisition: cracking the speech code. Nat. Rev. Neurosci. 5, 831–843. https://doi.org/10.1038/nrn1533.

Kuhl, P.K., Tsao, F.-M., Liu, H.-M., 2003. Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. Proc. Natl. Acad. Sci. Unit. States Am. 100, 9096–9101.

Kujala, A., Alho, K., Service, E., Ilmoniemi, R.J., Connolly, J.F., 2004. Activation in the anterior left auditory cortex associated with phonological analysis of speech input: localization of the phonological mismatch negativity response with MEG. Cognit. Brain Res. 21, 106–113. https://doi.org/10.1016/j.cogbrainres.2004.05.011.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.

Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cognit. Sci. 4, 463–470.

Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.

Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.

Lam, N.H.L., Schoffelen, J.-M., Uddén, J., Hultén, A., Hagoort, P., 2016. Neural activity during sentence processing as reflected in theta, alpha, beta, and gamma oscillations. Neuroimage. https://doi.org/10.1016/j.neuroimage.2016.03.007.

Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (de)constructing the N400. Nat. Rev. Neurosci. 9, 920–933. https://doi.org/10.1038/nrn2532.

Lee, T.-W., Girolami, M., Sejnowski, T.J., 1999. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. Neural Comput. 11, 417–441.

Lewis, A.G., Bastiaansen, M., 2015. A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. Cortex 68, 155–168. https://doi.org/10.1016/j.cortex.2015.02.014.

Lewis, A.G., Lemhöfer, K., Schoffelen, J.-M., Schriefers, H., 2016. Gender agreement violations modulate beta oscillatory dynamics during sentence comprehension: a comparison of second language learners and native speakers. Neuropsychologia 89, 254–272. https://doi.org/10.1016/j.neuropsychologia.2016.06.031.

Lewkowicz, D.J., Hansen-Tift, A.M., 2012. Infants deploy selective attention to the mouth of a talking face when learning speech. Proc. Natl. Acad. Sci. Unit. States Am. 109, 1431–1436. https://doi.org/10.1073/pnas.1114783109.

Li, P., Grant, A., 2016. Second language learning success revealed by brain networks. Biling. Lang. Cognit. 19, 657–664. https://doi.org/10.1017/S1366728915000280.

Liang, L., Chen, B., 2019. The impact of language proficiency on the time course and neural basis of L2 semantic access in bilinguals. Int. J. BiLing. 136700691989785 https://doi.org/10.1177/1367006919897851.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., Studdert-Kennedy, M., 1967. Perception of the speech code. Psychol. Rev. 74, 431–461. https://doi.org/10.1037/h0020279.

Luo, Y., Zhang, Y., Feng, X., Zhou, X., 2010. Electroencephalogram oscillations differentiate semantic and prosodic processes during sentence reading. Neuroscience 169, 654–664. https://doi.org/10.1016/j.neuroscience.2010.05.032.

Macedonia, M., 2013. Pronunciation in foreign language: how to train? Journal of Education and Training Studies 2. https://doi.org/10.11114/jets.v2i1.205.

Mårtensson, J., Eriksson, J., Bodammer, N.C., Lindgren, M., Johansson, M., Nyberg, L., Lövdén, M., 2012. Growth of language-related brain areas after foreign language learning. Neuroimage 63, 240–244. https://doi.org/10.1016/j.neuroimage.2012.06.043.

Mechelli, A., Crinion, J.T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R.S., Price, C.J., 2004. Neurolinguistics: structural plasticity in the bilingual brain. Nature 431, 757–757.

Meisel, J.M., 2011. First and Second Language Acquisition: Parallels and Differences. Cambridge University Press.

Menning, H., Imaizumi, S., Zwitserlood, P., Pantev, C., 2002. Plasticity of the human auditory cortex induced by discrimination learning of non-native, mora-timed contrasts of the Japanese language. Learn. Mem. 9, 253–267. https://doi.org/10.1101/lm.49402.

Midgley, K.J., Holcomb, P.J., Grainger, J., 2009. Language effects in second language learners and proficient bilinguals investigated with event-related potentials. J. Neurolinguistics 22, 281–300. https://doi.org/10.1016/j.jneuroling.2008.08.001.

Mitchell, R., Myles, F., Marsden, E., 2019. Second Language Learning Theories. Routledge.

Mittag, M., Takegata, R., Winkler, I., 2016. Transitional probabilities are prioritized over stimulus/pattern probabilities in auditory deviance detection: memory basis for predictive sound processing. J. Neurosci. 36, 9572–9579. https://doi.org/10.1523/JNEUROSCI.1041-16.2016.

Mognon, A., Jovicich, J., Bruzzone, L., Buiatti, M., 2011. ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features: automatic spatio-temporal EEG artifact detection. Psychophysiology 48, 229–240. https://doi.org/10.1111/j.1469-8986.2010.01061.x.

Morgan-Short, K., 2014. Electrophysiological approaches to understanding second language acquisition: a field reaching its potential. Annu. Rev. Appl. Ling. 34, 15–36. https://doi.org/10.1017/S026719051400004X.

Morgan-Short, K., Steinhauer, K., Sanz, C., Ullman, M.T., 2012. Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. J. Cognit. Neurosci. 24, 933–947. https://doi.org/10.1162/jocn_a_00119.

Näätänen, R., Gaillard, A.W., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. Acta Psychol. 42, 313–329.

Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. Clin. Neurophysiol. 118, 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026.

Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24, 375–425.

Näätänen, R., Winkler, I., 1999. The concept of auditory stimulus representation in cognitive neuroscience. Psychol. Bull. 125, 826–859. https://doi.org/10.1037/0033-2909.125.6.826.

Newman, A.J., Tremblay, A., Nichols, E.S., Neville, H.J., Ullman, M.T., 2012. The influence of language proficiency on lexical semantic processing in native and late learners of English. J. Cognit. Neurosci. 24, 1205–1223. https://doi.org/10.1162/jocn_a_00143.

Newman, R.L., Connolly, J.F., 2009. Electrophysiological markers of pre-lexical speech processing: evidence for bottom–up and top–down effects on spoken word processing. Biol. Psychol. 80, 114–121. https://doi.org/10.1016/j.biopsycho.2008.04.008.

Newport, E.L., Bavelier, D., Neville, H.J., 2001. Critical thinking about critical periods: perspectives on a critical period for language acquisition. Language, brain and cognitive development. Essays in honor of Jacques Mehler 481–502.

Obleser, J., Elbert, T., Lahiri, A., Eulitz, C., 2003. Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. Cognit. Brain Res. 15, 207–213. https://doi.org/10.1016/S0926-6410(02)00193-3.

Obleser, J., Kotz, S.A., 2011. Multiple brain signatures of integration in the comprehension of degraded speech. Neuroimage 55, 713–723. https://doi.org/10.1016/j.neuroimage.2010.12.020.

Omote, A., Jasmin, K., Tierney, A., 2017. Successful non-native speech perception is linked to frequency following response phase consistency. Cortex 93, 146–154. https://doi.org/10.1016/j.cortex.2017.05.005.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 1–9. https://doi.org/10.1155/2011/156869.

Ouyang, G., Sommer, W., Zhou, C., 2015a. A toolbox for residue iteration decomposition (RIDE)—a method for the decomposition, reconstruction, and single trial analysis of event related potentials. J. Neurosci. Methods 250, 7–21. https://doi.org/10.1016/j.jneumeth.2014.10.009.

Ouyang, G., Sommer, W., Zhou, C., 2015b. Updating and validating a new framework for restoring and analyzing latency-variable ERP components from single trials with residue iteration decomposition (RIDE): ERP analysis with residue iteration decomposition. Psychophysiology 52, 839–856. https://doi.org/10.1111/psyp.12411.

Pallier, C., Bosch, L., Sebastián-Gallés, N., 1997. A limit on behavioral plasticity in speech perception. Cognition 64, B9–B17. https://doi.org/10.1016/S0010-0277(97)00030-9.

Penolazzi, B., Angrilli, A., Job, R., 2009. Gamma EEG activity induced by semantic violation during sentence reading. Neurosci. Lett. 465, 74–78. https://doi.org/10.1016/j.neulet.2009.08.065.

Pérez, A., Carreiras, M., Gillon Dowens, M., Duñabeitia, J.A., 2015. Differential oscillatory encoding of foreign speech. Brain Lang. 147, 51–57. https://doi.org/10.1016/j.bandl.2015.05.008.

Picard, M., 2002. The differential substitution of English/θ ð/in French: the case against underspecification in L2 phonology. Lingvisticae Investigationes 25, 87–96. https://doi.org/10.1075/li.25.1.07pic.

Polivanov, E., 1931. La perception des sons d'une langue étrangère. Travaux du Cercle Linguistique de Prague 4, 79–96.

Pruitt, J.S., Jenkins, J.J., Strange, W., 2006. Training the perception of Hindi dental and retroflex stops by native speakers of American English and Japanese. J. Acoust. Soc. Am. 119, 1684–1696. https://doi.org/10.1121/1.2161427.

Pulvermüller, F., Shtyrov, Y., 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. Prog. Neurobiol. 79, 49–71. https://doi.org/10.1016/j.pneurobio.2006.04.004.

R Core Team, 2013. R: A Language and Environment for Statistical Computing.

Reetzke, R., Xie, Z., Llanos, F., Chandrasekaran, B., 2018. Tracing the trajectory of sensory plasticity across different stages of speech learning in adulthood. Curr. Biol. 28 (e4), 1419–1427. https://doi.org/10.1016/j.cub.2018.03.026.

Rihs, T.A., Michel, C.M., Thut, G., 2007. Mechanisms of selective inhibition in visual spatial attention are indexed by α-band EEG synchronization. Eur. J. Neurosci. 25, 603–610. https://doi.org/10.1111/j.1460-9568.2007.05278.x.

Sanders, L.D., Newport, E.L., Neville, H.J., 2002. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. Nat. Neurosci. 5, 700–703. https://doi.org/10.1038/nn873.

Scharinger, M., Monahan, P.J., Idsardi, W.J., 2016. Linguistic category structure influences early auditory processing: converging evidence from mismatch responses and cortical oscillations. Neuroimage 128, 293–301. https://doi.org/10.1016/j.neuroimage.2016.01.003.

Schlegel, A.A., Rudelson, J.J., Peter, U.T., 2012. White matter structure changes as adults learn a second language. J. Cognit. Neurosci. 24, 1664–1670.

Sebastian-Galles, N., 2006. Native-language sensitivities: evolution in the first year of life. Trends Cognit. Sci. 10, 239–241. https://doi.org/10.1016/j.tics.2006.04.009.

Sebastián-Gallés, N., Echeverría, S., Bosch, L., 2005. The influence of initial exposure on lexical representation: comparing early and simultaneous bilinguals. J. Mem. Lang. 52, 240–255. https://doi.org/10.1016/j.jml.2004.11.001.

Sinclair, J.M., 1987. Looking up: an Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary. Collins Elt.

Stein, M., Federspiel, A., Koenig, T., Wirth, M., Strik, W., Wiest, R., Brandeis, D., Dierks, T., 2012. Structural plasticity in the language system related to increased second language proficiency. Cortex 48, 458–465. https://doi.org/10.1016/j.cortex.2010.10.007.

Terporten, R., Schoffelen, J.-M., Dai, B., Hagoort, P., Kösem, A., 2019. The relation between alpha/beta oscillations and the encoding of sentence induced contextual information. Sci. Rep. 9, 20255. https://doi.org/10.1038/s41598-019-56600-x.

Thornton, A.R.D., Harmer, M., Lavoie, B.A., 2007. Selective attention increases the temporal precision of the auditory N100 event-related potential. Hear. Res. 230, 73–79. https://doi.org/10.1016/j.heares.2007.04.004.

Troubetzkoy, N.S., 1939. Principes de phonologie. Klincksieck, Paris.

Tyler, M.D., 2021. Phonetic and phonological influences on the discrimination of non-native phones. In: Wayland, R. (Ed.), Second Language Speech Learning: Theoretical and Empirical Progress. Cambridge, UK, pp. 157–174.

Ullman, M.T., 2020. The declarative/procedural model: a neurobiologically-motivated theory of first and second language. In: VanPatten, B., Keating, G.D., Wulff, S. (Eds.), Theories in Second Language Acquisition. Routledge, pp. 128–161.

Van Berkum, J.J.A., 2009. The neuropragmatics of 'simple' utterance comprehension: an ERP review. In: Sauerland, U., Yatsushiro, K. (Eds.), Semantics and Pragmatics: from Experiment to Theory. Palgrave Macmillan, Basingstoke, pp. 276–316.

Van Petten, C., Kutas, M., 1990. Interactions between sentence context and word frequency in event-related brain potentials. Mem. Cognit. 18, 380–393. https://doi.org/10.3758/BF03197127.

Ventureyra, V.A.G., Pallier, C., Yoo, H.-Y., 2004. The loss of first language phonetic perception in adopted Koreans. J. Neurolinguistics 17, 79–91. https://doi.org/10.1016/S0911-6044(03)00053-8.

Veroude, K., Norris, D.G., Shumskaya, E., Gullberg, M., Indefrey, P., 2010. Functional connectivity between brain regions involved in learning words of a new language. Brain Lang. 113, 21–27. https://doi.org/10.1016/j.bandl.2009.12.005.

Wang, L., Hagoort, P., Jensen, O., 2018. Gamma oscillatory activity related to language prediction. J. Cognit. Neurosci. 30, 1075–1085. https://doi.org/10.1162/jocn_a_01275.

Wang, L., Jensen, O., van den Brink, D., Weder, N., Schoffelen, J.M., Magyari, L., Hagoort, P., Bastiaansen, M., 2012. Beta oscillations relate to the N400m during language comprehension. Hum. Brain Mapp. 33, 2898–2912. https://doi.org/10.1002/hbm.21410.

Weber-Fox, C.M., Neville, H.J., 1996. Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. J. Cognit. Neurosci. 8, 231–256. https://doi.org/10.1162/jocn.1996.8.3.231.

Werker, J.F., Tees, R.C., 2005. Speech perception as a window for understanding plasticity and commitment in language systems of the brain. Dev. Psychobiol. 46, 233–251. https://doi.org/10.1002/dev.20060.

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

White, E.J., Hutka, S.A., Williams, L.J., Moreno, S., 2013. Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. Front. Syst. Neurosci. 7 https://doi.org/10.3389/fnsys.2013.00090.

Wilkinson, M.D., Dumontier, M., Aalbersberg, IjJ., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018. https://doi.org/10.1038/sdata.2016.18.

Xiang, H., van Leeuwen, T.M., Dediu, D., Roberts, L., Norris, D.G., Hagoort, P., 2015. L2-Proficiency-Dependent laterality shift in structural connectivity of brain language pathways. Brain Connect. 5, 349–361. https://doi.org/10.1089/brain.2013.0199.