Bilingualism: Language and Cognition

cambridge.org/bil

The neurophysiology of phonemic contrasts perception in L2/L3 learners: The role of acquisition setting

Hanna Kędzierska[1,2], Karolina Rataj[3], Anna Balas[1] and Magdalena Wrembel[1]

[1]Department of Contemporary English and Multilingualism, Faculty of English, Adam Mickiewicz University, Poznań, Poland; [2]Department of English and Comparative Linguistics, Institute of English Studies, University of Wrocław, Wrocław, Poland and [3]Neuroscience of Language Laboratory, Faculty of English, Adam Mickiewicz University, Poznań, Poland

Research Article

Cite this article: Kędzierska, H., Rataj, K., Balas, A. and Wrembel, M. (2025). The neurophysiology of phonemic contrasts perception in L2/L3 learners: The role of acquisition setting. *Bilingualism: Language and Cognition* **28**, 550–562. https://doi.org/10.1017/S136672892400066X

Received: 3 January 2024
Revised: 6 July 2024
Accepted: 6 July 2024
First published online: 27 November 2024

Keywords:
multilingualism; phonology; third language (L3/Ln); mismatch negativity (MMN); language acquisition

Corresponding author:
Hanna Kędzierska;
Email: hanna.kedzierska@uwr.edu.pl

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

Abstract

Phoneme discrimination is believed to be less accurate in non-native languages compared to native ones. What remains unclear is whether differences in pre-attentive phonological processing emerge between the first foreign language (L2) and additional ones (L3/L*n*), and whether they might be influenced by the acquisition setting (formal vs. naturalistic). We conducted an event-related brain potential oddball study with native Polish learners of English (L2) and Norwegian (L3/L*n*). The results revealed a graded amplitude of the mismatch negativity (MMN) effect, which was largest in L1, smaller in L2, and smallest in L3/L*n*. Considering the previously obtained results for naturalistic/mixed learners with the same language combination, we believe that the acquisition setting is an important factor influencing the perception of phonemic contrasts. In the naturalistic group, no difference was observed between L1 and L2, while the instructed group exhibited more fine-grained distinctions between all tested languages.

1. Introduction

The focus of this paper is on non-native phoneme discrimination in multilingual learners, i.e., those who, in addition to their second language (L2), have acquired at least one additional foreign language (L3/L*n*). The field of L3/L*n* phonology has been extensively explored in recent years, with the last decade witnessing manifold studies on the phonological system developed by multilinguals. Notably, the majority of this research has focused on speech production rather than perception, although this trend seems to be partially reversing in recent years (see Wrembel, 2023, for a review). Still, a modest number of researchers have directly explored the processing of non-native phonemes in L2 when compared with L3/L*n*, and the experimental techniques used to investigate the non-native phoneme perception in multilinguals have typically involved forced-choice goodness tasks (e.g., Cabrelli Amaro, 2017), perceptual assimilation tasks (e.g., Kopečková, 2015), foreign-accentedness ratings (e.g., Lloyd-Smith, 2023; Wrembel et al., 2019), or discrimination tasks (e.g., Wrembel et al., 2019). What has been scarcely studied so far is the neural markers of phonemic contrast perception by multilingual listeners, with most of the work prioritizing bilingual research (e.g., Jakoby et al., 2011; Liang & Chen, 2022; Song & Iverson, 2018). In order to bridge this gap, we report the results of a phonemic perception event-related potential (ERP) study, which focused on trilingual L1 Polish – L2 English – L3 Norwegian speakers who acquired their L3/L*n* in a formal (i.e., classroom-instruction) setting. The selection of trilingual instead of bilingual listeners enabled us to elucidate the issue of language status (operationalized as L2 or L3/L*n*) as a factor affecting non-native phoneme processing.

Categorical perception (Liberman et al., 1957) is an empirical phenomenon, when the ability to distinguish two sounds is affected by the categories they belong to. In a native language, we can discriminate sounds that belong to two different phonemes, but not when they belong to a single category. Categorical perception helps us process the incoming speech signal fast in the L1, but in a non-native language, it tends to hinder differentiation between phones that belong to two different phonemes in L2, but are assimilated to a single L1 phoneme. Major theories of L2 acquisition of speech, such as the Speech Learning Model (Flege, 1995), revised Speech Learning Model (Flege & Bohn, 2021), Perceptual Assimilation Model – L2 (Best & Tyler, 2007) and Second Language Perception Model (Escudero, 2005, 2009; Escudero & Boersma, 2004) associate category formation in a non-native language with the ability to detect differences between L1 and target language sounds. Whereas a large body of research dealt with studies of conscious non-native phone discrimination using behavioral techniques (e.g., Flege et al., 1994; Best et al., 2001; Levy & Strange, 2008; Tyler et al., 2014) the present study concentrates on neural, pre-attentional

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S136672892400066X

discrimination of native and non-native sounds. This is particularly novel with respect to investigating online phonemic perception in multilinguals, comparing two non-native languages.

As a next step, we juxtapose the results of the current study with the data obtained earlier in the same task by a group of trilingual speakers who acquired the foreign languages (i.e., English and Norwegian) in a naturalistic (i.e., immersion) setting (Kędzierska et al., 2023). In this way, we examine whether the learning context (i.e., the acquisition setting) may potentially affect the neural discrimination of phonemic contrasts. Even more crucially, the current research also contributes to the ongoing debate concerning the effectiveness of foreign language learning, with non-native phoneme perception and consequently, their production as its central components.

One of the most pertinent questions both within the field of L2 and L3/L*n* acquisition concerns the effectiveness of the learning process shaped by the learning context (or setting), with the nature of this context typically instantiated as either naturalistic (i.e., immersive) or formal (i.e., instructed). The former one typically progresses in natural communicative contexts, holistically, without interruptions and in a country or region where the target language is used on an everyday basis. The latter one is assumed to be "notably less than ideal with respect to the natural ecology of language learning" (Best & Tyler, 2007, p. 19), as it takes place in a restricted classroom setting, does not extend much outside classroom time, it happens in smaller portions, often with the learner's L1 used for instruction, and usually centered around grammatical and lexical information. Moreover, the exposure quality in the formal classroom setting is specific, because it primarily involves interactions with L1-accented teacher and fellow classmates, entailing many mistargeted phonetic details. Obviously, there are also numerous other acquisition contexts situated somewhere in between these two archetypal learning scenarios (see Muñoz, 2008, for a discussion). One could mention here, for instance, a mixed context where formal education proceeds by means of immersion or self-learning attempts facilitated by such resources as podcasts, videos, or language learning apps. However, since most studies to date have focused on non-native language acquisition in either immersive or formal settings, an apparent need arises for more systematic comparisons of patterns observed among two learner types with shared language repertoires (e.g., Muñoz, 2008; Puig-Mayenco et al., 2020; Wrembel, 2023). Hence, in the current paper, we will focus on two markedly different learning scenarios, i.e., one involving formal classroom instructions of L2 English and L3/L*n* Norwegian (the study reported in this paper) as compared to the one involving naturalistic acquisition of L3/L*n* Norwegian and naturalistic/mixed[1] acquisition of L2 English by Polish migrants in Norway (Kędzierska et al., 2023).

2. Study motivation and background

The method employed in the current research is event-related brain potentials (ERP), which, thanks to its high temporal resolution, is perfectly suited to track such rapidly occurring cognitive processes as phoneme recognition (Kaan, 2007). These processes are reflected in the ERP components, i.e., parts of the ERP waveforms elicited as a reaction to the manipulation (Woodman, 2010, p. 4). The component that is typically associated with the listeners' sensitivity to discriminable phonemic changes presented within a series of sounds is the mismatch negativity (MMN) (Kujala & Näätänen, 2003; Näätänen et al., 1997, 2007). This component is typically elicited in oddball tasks, in which a series of standard stimuli (e.g., identical sounds) is occasionally interrupted by the occurrence of a deviant stimulus (e.g., a different sound). The occurrence of the deviant violates the prediction that the listener has made on the basis of standard sounds. As a consequence, the MMN effect – defined as the difference in voltage between the deviant and standard sound – is elicited over fronto-central sites at around 150–250 milliseconds (ms) from the onset of the deviant stimulus (Näätänen et al., 2007). Quite importantly, this brain response is believed to be elicited pre-attentively, and hence does not require the participants' attention, which is typically focused on an unrelated primary task, such as watching a video or reading a book. The MMN is often followed by late discriminative negativity (LDN), i.e., a negativity (also observed over fronto-central sites) at around 350–600 ms after the onset of the deviant. While the functional role of the LDN has not yet been fully established (see Jakoby et al., 2011, for a discussion), the component is typically associated with pre-attentive cognitive evaluation of the stimulus (Ceponiene et al., 1998; Liang & Chen, 2022) or a successful formation of memory traces associated with a specific phonemic representation (Barry et al., 2009; Jakoby et al., 2011). In general, in the context of non-native phoneme processing, previous studies observed reduced MMN in non-native languages when compared with L1 (Jakoby et al., 2011; Liang & Chen, 2022; Song & Iverson, 2018) as well as in L3/L*n* when compared with L2 (Kędzierska et al., 2023). At the same time, the LDN effect, which followed the MMN in some ERP studies on bilingual phoneme processing, was typically higher in more successful/higher proficiency L2 learners as opposed to less successful/lower proficiency ones (Jakoby et al., 2011 and Liang & Chen, 2022, respectively). In Kędzierska et al. (2023), no significant difference was found between L2 and L3/L*n* in terms of the LDN amplitude, but for both foreign languages, the effect was smaller when compared with L1.

Notably, research on bilingual listeners' phonological sensitivity has also produced several conflicting findings, suggesting the significance of listener-oriented factors in the processing of non-native phonemic contrasts. One of these factors seems to involve the learning setting differences discussed above. In the study conducted by Winkler et al. (1999), native speakers of Finnish and late learners of Finnish in an immersive context both showed a similar MMN effect to Finnish vowel contrasts. On the other hand, Peltola et al. (2003) discovered a significant difference between native English speakers and instructed learners: for the latter ones, English vowel contrasts elicited lower MMN amplitudes (a finding also supported by Wottawa et al., 2022). In a recent study, Kędzierska et al. (2023) examined how sensitive the MMN effect is to phoneme contrasts in L1 Polish, L2 English (naturalistic/mixed), and L3/L*n* Norwegian (naturalistic) learners. Language status (defined in terms of dominance rather than chronology) had an impact on the MMN effect, which was reflected in significant differences between L1 and L3/L*n* as well as between L2 and L3/L*n*. More specifically, the amplitude of the MMN effect was smaller for L3/L*n* when compared with both L1 and L2.

In the current study, we aimed to contribute to an ongoing scientific debate on multilingual phoneme processing by investigating the perception of L1 Polish, L2 English, and L3/L*n* Norwegian vowels among formal L2 English and L3/L*n* Norwegian learners. In order to determine whether the listeners' phonological systems are sensitive to the non-native vowel contrasts, we decided

[1]In this study, the vast majority of the participants started learning English at school before the puberty age and then in the adulthood migrated to Norway where they predominantly used English in the work environment. Hance, we describe them as naturalistic/mixed learners.

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/S136672892400066X

to use the MMN as a neurophysiological marker reflecting ease of phonemic discrimination (Näätänen et al., 1997). Therefore, an ERP experiment was designed with a passive oddball paradigm, where the vowels in each language were presented independently in three blocks.

To control the conditions well, it would be desirable to have the same standard in all three languages and deviants that are equidistant and are not different in terms of phonetic features, so that we could examine the effect of language status on their processing. Such ideal conditions, however, cannot be found in natural languages, so we need to study various scenarios, as each of them contributes to a bigger picture encompassing numerous variables that influence acquisition of speech in non-native languages. In the present paper, we used a common standard in all three languages and different, non-overlapping deviants. The deviants differed in terms of phonetic features and also the Euclidean distances between the standard and the deviant varied across languages, which was taken into consideration in the analysis and interpretation of the results. In order to account for the acoustic properties of vowels, we also analyzed formant distances between vowels in a given language pair (see Table S1 in Appendix S1), as well as between language pairs (see Table S2 in Appendix S1), to mirror the analysis conducted on MMN differences and see which phonetic component is most closely related to the MMN differences.

The design we employed permitted three research questions to be addressed and several predictions to be made. First of all, we were interested in whether we would observe any significant difference in phoneme processing as indexed by the MMN and LDN in the native language (i.e., L1 Polish), when compared with non-native languages (i.e., L2 English and/or L3/L*n* Norwegian). Since previous studies in bilingual phoneme processing indicate that the MMN effect is typically smaller in the non-native when compared with native languages, especially if acquired in a classroom setting (e.g., Jakoby et al., 2011; Liang & Chen, 2022; Song & Iverson, 2018), we expected that the effect would be significantly larger in L1 Polish when compared with L2 English and L3/L*n* Norwegian. At the same time, based on the research of Kędzierska et al. (2023), we expected the LDN effect to be significantly smaller in non-native languages when compared with L1.

Further, we wished to answer a research question concerning any significant differences in the MMN or LDN in L3/L*n* Norwegian as opposed to L2 English. On the basis of previous L2 research (e.g., Jakoby et al., 2011; Liang & Chen, 2022; Song & Iverson, 2018) as well as research involving multilingual listeners (Kędzierska et al., 2023), we could tentatively assume that the MMN effect in L3/L*n* (quite crucially, defined in terms of proficiency and dominance, rather than chronology) would be smaller relative to L1, and similar or smaller relative to L2. We also predicted that the effect might be influenced by the age of acquisition (AoA) and/or proficiency in non-native languages. However, again, based on the findings of Kędzierska et al. (2023), we did not predict any significant differences in terms of the LDN component when comparing L2 English and L3/L*n* Norwegian.

Finally, we also wished to explore the potential neural differences expressed in terms of the MMN and LDN effects for the L3/L*n* Norwegian contrast between formal/instructed-setting learners as opposed to naturalistic language learners. We expected the MMN effect in L3 to be enhanced in naturalistic when compared with instructed language learners due to more exposure to L3 Norwegian (Peltola et al., 2003; Winkler et al., 1999). To enable such a comparison, we decided to analyze the current results in light of the ones obtained earlier by Kędzierska et al. (2023). Due to the lack of previous research that would focus specifically on the influence of language setting on the LDN, no specific predictions could be made concerning the LDN effect.

3. Methods

3.1. Participants

Twenty-two participants (mean age = 22.55, age range: 18–38, 17 females, five males) were recruited to take part in the study. Twenty participants were right-handed and two were mixed-handed[2] as assessed by the Edinburgh Handedness Inventory (adapted from Oldfield, 1971), with the mean laterality quotient (LQ) equal to 85.91% (range: 45.00%–100.00%, *SD* = 16.74%). All of the participants were college students (*N* = 18) or college graduates (*N* = 4) whose native language was Polish, and who also spoke English (L2) and Norwegian (L3/Ln). For all the participants, Polish was their only native language, and English was the first foreign language they started learning at school or preschool before puberty. For the majority of the participants, Norwegian was chronologically either the third (*N* = 9) or the fourth (*N* = 11) language; however, there were also two participants for whom it was chronologically the fifth and the sixth language. The age of acquisition, according to self-report, was 5.86 years (range: 3–10, *SD* = 1.83) for L2 English and 20.27 years (range: 13–36, *SD* = 4.73) for L3/L*n* Norwegian. At the time of the experiment, 18 participants were enrolled in the BA (*N* = 16) or MA (*N* = 2) program of Scandinavian Studies (i.e., Norwegian philology) at the Adam Mickiewicz University or the College of Modern Languages in Poznań, Poland, and two had recently graduated from one of these programs (i.e., a year earlier, *N* = 2). Additionally, two participants graduated from other courses (i.e., geography and civil engineering) and learnt Norwegian at a private language school. None of the participants reported any neurological and psychiatric impairments nor any language-related issues (such as dyslexia or dysorthography). All participants received gift cards as compensation for their time.

The participants were asked to self-assess their knowledge of Polish, English, and Norwegian in listening, speaking, reading and writing skills on a scale from 1 (very low) to 7 (proficient). They all assessed their Polish skills as 7.00 (range: 7.00–7.00, *SD* = 0), their English skills as 5.85 on average (range: 5.25–7, *SD* = 0.60), and their Norwegian skills as 3.88 on average (range: 2–6.5, *SD* = 1.16). In addition, the participants' knowledge of the two foreign languages was verified with the aid of the Cambridge General English Assessment Test and the UiT Norwegian Placement Test, both taken immediately after the EEG session in the presence of the experimenters. The average result of the English proficiency test equaled 82.00% (range: 40.00%–100.00%, *SD* = 16.22%) and that of the Norwegian proficiency test equaled 65.78% (range: 22.22%–91.67%, *SD* = 22.41%). A more detailed summary of the participants' biographic data and proficiency results is included in Table 1.

3.2. Stimuli

In order to select the sounds that would be used as stimuli in the current study, a comparison of phonological systems of Polish, English and Norwegian was undertaken. The decision to

[2]The statistical analyses produced consistent results whether or not the data from the two mixed-handed participants were included. Therefore, their data was retained in the final analyses.

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/S136672892400066X

Table 1. The summary of the participants' biographic details and language proficiency

Participants' profiles	
Biographic details	
Age	*M* = 22.5, range: 18–38, *SD* = 4.5
Gender	17 females, 5 males
Other languages	German (8), Russian (3), Spanish (2), Swedish (2), Icelandic (1), Czech (1), Danish (1), Chinese (1), Korean (1), Croatian (1), Ukrainian (1)
Proficiency self-assessment	
L1 Polish	*M* = 7.00, range: 7–7, *SD* = 0.00
L2 English (overall)	*M* = 5.85, range: 4–7, *SD* = 0.60
L2 English (listening)	*M* = 5.82, range: 4–7, *SD* = 0.73
L2 English (speaking)	*M* = 5.64, range: 4–7, *SD* = 0.79
L2 English (reading)	*M* = 6.09, range: 5–7, *SD* = 0.61
L2 English (writing)	*M* = 5.86, range: 4–7, *SD* = 0.77
L3/L*n* Norwegian (overall)	*M* = 3.88, range: 2–6.5, *SD* = 1.16
L3/L*n* Norwegian (listening)	*M* = 3.73, range: 2–7, *SD* = 1.32
L3/L*n* Norwegian (speaking)	*M* = 3.59, range: 2–6, *SD* = 1.14
L3/L*n* Norwegian (reading)	*M* = 4.32, range: 2–7, *SD* = 1.25
L3/L*n* Norwegian (writing)	*M* = 3,86, range: 2–6, *SD* = 1.17
Proficiency tests results	
L2 English	*M* = 82.00%, range: 40.00%–100.00%, *SD* = 16.22%
L3/L*n* Norwegian	*M* = 65.78%, range: 22.22%–91.67%, *SD* = 22.41%

use vowels as experimental stimuli was motivated by previous electrophysiological research on non-native phoneme sensitivity, where vowels were used fairly frequently, either in isolation (e.g., Díaz et al., 2016; Liang & Chen, 2022; Peltola et al., 2012; Winkler et al., 1999), or embedded in syllables (e.g., Jakoby et al., 2011; White et al., 2017). We decided against using syllables as stimuli because the consonants would carry over additional language-related differences, and we would not be sure what drives listeners' reactions – differences in vowels or consonants. Moreover, the three languages used in the current study differ considerably with respect to their vowel inventory density, making vowel perception and production particularly problematic for foreign language learners. While Polish has a rather small vowel inventory, with only six monophthongal vowels (Jassem, 2003), the vocalic systems of English and Norwegian are richer with 12 and 18 monophthongal vowels, respectively (Wells, 1962; Upton et al., 2003; Bjelaković, 2016; Kristoffersen, 2000). Furthermore, the three languages in question differ with respect to the combination of lip-rounding with backness: while all of them have front unrounded vowels and back rounded vowels, English and Norwegian have high central rounded vowels, but only Norwegian has front rounded vowels – a configuration which seems generally dispreferred (or marked) among world languages (Maddieson, 2013).

As the participants of the current study acquired Polish since birth, English in childhood, and Norwegian in adulthood, in their case, the order of acquisition would presume a gradual enlargement of phonemic (and, specifically, vocalic) repertoire. Consequently, our choice of standard stimuli was motivated by the high degree of cross-linguistic similarity between the three standard sounds, i.e., Polish /i/, English /ɪ/ and Norwegian /i/. On the other hand, the selection of deviants was motivated by systematic differences between the three languages studied, with an additional criterion that the standard and deviant sounds in each language should be comparably distanced from one another. Thus, we selected the following pairs: /i/–/ɛ/ for Polish, /ɪ/–/ʊ/ for English, and /i/–/ʏ/ for Norwegian. The Polish /i/–/ɛ/ contrast is expressed mainly in vowel height and also exists in the other languages under investigation. The English /ɪ/–/ʊ/ contrast manifests itself mainly in backness and rounding, and is also present in Norwegian, but not in Polish, a language with no near-high rounded vowels. Finally, the /i/–/ʏ/ contrast in Norwegian is expressed mainly through roundness and is absent in Polish and English, which do not have front-rounded vowels.

In order to address the research questions and to compare the results of the current study directly to those obtained earlier by Kędzierska et al. (2023) for naturalistic language learners, we used the same experimental stimuli involving a set of six monophthongal vowels: /i/ and /ɛ/ for Polish (as in the Polish words *byty* 'being$_{PL}$' and *bety* 'bed linen$_{PL}$'), /ɪ/ and /ʊ/ for English (as in *fit* and *foot* respectively) and /i/ and /ʏ/ for Norwegian (as in the Norwegian words *sin* 'his$_{REFL}$' and *synd* 'shame' respectively). The vowels used in the current study were synthesized with the aid of the PRAAT software (Boersma, 2001) based on formant frequency values available in the literature (Weckwerth & Balas, 2019 for Polish; Bjelaković, 2017 for English). Due to the lack of similar reference data for Norwegian, the Norwegian vowels were generated based on the average values obtained from four male native speakers of Norwegian. For all the synthesized stimuli, the duration was 150 ms, the amplitude contour had a 3 ms linear onramp and 75 ms linear offramp, and the f0 trajectory had a steady linear fall from 140 Hz to 110 Hz. The formant values for each vowel as well as Euclidean distances between vowels presented in the same language pairs are presented in Table 2. We also report formant distances between vowels in a given language pair in Appendix S1 (Table S1) to analyze the distances between language pairs (see Table S2) and to establish which phonetic parameters are closely related to the differences reported for MMN.

3.3. Procedures

The participants were comfortably seated in a sound-attenuated room. At the beginning of the experimental session, they were asked to complete a language history questionnaire (LHQ) (based on Li et al., 2020) and the Edinburgh Handedness Inventory (adapted from Oldfield, 1971). Further, the experimental stimuli were presented over earphones in three language blocks, the order of which was counterbalanced across participants. Each block contained 660 trials, including 600 standard trials and 60 deviant trials, which means that the probability of a deviant sound occurrence equaled 9.00% in each experimental block. The order of trials was pseudorandomized (independently for every participant), so that each deviant sound was preceded by at least three standard sounds. The volume level was kept constant. While the stimuli were presented, the participants watched a silent movie (more specifically, a cartoon titled *Bolek i Lolek*). They were explicitly instructed (in Polish) to focus their attention on the cartoon rather than on the sounds and informed that they would be asked 10 detailed open questions about the

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/S136672892400066X

Table 2. The summary of vowel formant frequencies used for stimuli synthesis (in Hz) and Euclidean distances between vowels (in Hz and Bark)

Vowel	F1 (Hz)	F1 (Bark)	F2 (Hz)	F2 (Bark)	F3 (Hz)	F3 (Bark)
Polish /i/	468	4.64	1948	12.83	2821	15.29
Polish /ɛ/	675	6.34	1916	12.72	2722	15.06
English /ɪ/	394	3.96	1828	12.41	2882	15.43
English /ʊ/	390	3.92	1345	10.38	2896	15.46
Norwegian /i/	357	3.60	1917	12.73	2587	14.72
Norwegian /ʏ/	313	3.16	2015	13.06	2708	15.02
Euclidean distance	**/i/: /ɛ/**	**/ɪ/: /ʊ/**	**/i/: /ʏ/**			
F1-F2 (Hz)	209	483	107			
F1-F2 (Bark)	2.05	4.77	1.06			
F1-F2-F3 (Hz)	232	483	161			
F1-F2-F3 (Bark)	2.31	4.77	1.51			

content of the movie at the end of the experimental session. The language blocks were separated by short breaks of approximately 3 min during which the participants continued watching the cartoon in silence. At the end of the experimental session, the participants completed several tests: a movie comprehension quiz, a gating task aimed at assessing their phonological aptitude (understood as a general ability to differentiate between sounds), as well as English and Norwegian multiple choice proficiency tests.

Gating task

In order to control for any potential individual differences, a gating task in English was administered to the participants to tap into speech-specific capabilities in a foreign language. English was selected as the language of the task, due to its being chronologically the first and more advanced foreign language of the participants. The gating task followed a two-alternative forced choice test procedure, as used by Sebastián-Gallés and Soto-Faraco (1999) and Sebastian-Gallés and Baus (2005). A similar test was also employed by Díaz et al. (2016) in order to assess bilingual participants as good or poor L2 perceivers (Díaz et al., 2016, p. 959). As in the study of Díaz et al. (2016), the participants were exposed to perceptual stimuli consisting of four minimal pairs with the /æ/−/ɛ/ contrast (i.e., BAG-BEG, LAUGHED-LEFT, SHALL-SHELL, GAS-GUESS). We chose this contrast to be used in the task because /æ/ is not part of the Polish vowel repertoire, and is hence one of the sounds the discrimination of which is problematic for native Polish speakers. The participants were asked to identify the word whose fragment was presented via earphones by pressing one of the two keys ("L" or "A") on the computer keyboard. Further, they were instructed to assess their confidence level on a 7-point Likert scale. The experimental tokens could be considered a continuum: each word was divided into 10 gates (i.e., fragments) by adding or subtracting 10 ms from the alineation point, i.e., the point where the token words from a given pair (e.g., BAG-BEG) started to diverge. The point was determined based on the visual inspection in PRAAT (Boersma, 2001). Each word token of the minimal pairs was presented twice, which totaled 160 trials (4 pairs × 2 words × 10 gates × 2 presentations), with a break after 80 trials. The words were recorded by a native speaker of American English and presented at an intensity of 75 dB. The gating experiment was conducted in PsychoPy (Peirce et al., 2019).

EEG recordings

The EEG signal was recorded using Brain Products actiCHamp acquisition device at a 500 Hz sampling rate from 64 active electrodes placed at the elastic cap according to the extended 10–20 convention. The ground was positioned at AFz. Two electrodes were placed at the outer canthus of each eye (HEOG1 and HEOG2) and two were placed below and above the right eye (VEOG1 and VEOG2). The signal was referenced online to the right mastoid bone (approximated from TP8), and later re-referenced offline to the average of right and left mastoid bones (approximated from TP7 and TP8). Electrode impedances were kept below 10 kΩ. The EEG data were processed with the Brain Vision Analyzer 2 software (Brain Products, Gilching).

EEG signal processing and data analysis

The pre-processing steps were the same as reported by Kędzierska et al. (2023): first, the data were filtered offline with a 0.1–30 Hz band-pass filter, followed by a semi-automatic ICA ocular correction and re-referencing. Epochs time-locked to the onset of each stimulus were extracted between −200 and 800 ms. All deviant stimuli ($N = 180$) and the standard stimuli immediately preceding a deviant stimulus ($N = 180$) were included in the analysis. Baseline correction was performed in reference to pre-stimulus activity (i.e., −200 to 0 ms). Epochs contaminated by ocular or muscular artifacts (maximal allowed voltage step: 50 μV/ms, maximal allowed difference of values in intervals of 200 ms: 200 μV, minimal allowed amplitude: −100 μV, maximal allowed amplitude: 100 μV) were rejected from further analysis. This resulted in the exclusion of 2.99% of trials (3.11% for Polish standards, 3.03% for Polish deviants, 3.18% for English standards, 2.50% for English deviants, 2.95% for Norwegian standards and 3.18% for Norwegian deviants).

For each participant, we first computed the averaged waveforms for the standard and the deviant stimuli. After that the difference waveforms (i.e., deviant minus standard) were created. Following Luck and Gaspelin (2017), we defined the time windows used further in the statistical analysis based on the so-called 'collapsed' waveforms (in our case, averaged waveforms elicited by all standard and deviant stimuli, ignoring the three language conditions). As in the case of Kędzierska et al. (2023), this approach revealed an increased negativity in the 100–200 ms time window, which was followed by a late negativity in the 350–800 ms time window.

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/S136672892400066X

Notably, the selected time windows have already been used in the literature: both in the case of MMN (Kujala & Näätänen, 2003) and LDN (Di Dona et al., 2022). We only included electrodes from the fronto-central brain area in the analysis (F1, F2, F3, F4, Fz, FC1, FC2, FC3, FC4, FCz, C1, C2, C3, C4, Cz), since both investigated effects are typically observed over these sites (Ceponiene et al., 1998; Kujala & Näätänen, 2003).

In order to compare the effect sizes for significant effects observed in the lme analysis, we calculated the difference wave (i.e., deviant minus standard) for each participant, individually in each of the analyzed electrodes. We used the lme4 package (Bates et al., 2012) in the R software (R Core Team, 2012) to perform a linear mixed effects analysis to assess the relationship between the investigated languages and the size of the MMN and LDN effects. The procedure was conducted twice: in the earlier time window (i.e., 100–200 ms) for the MMN effect[3] and in the later time window (i.e., 350–800 ms) for the LDN effect. Language (i.e., Polish, English and Norwegian) was included in the model as a fixed effect and the intercepts for participants and for electrodes were included as random effects. In the MMN analysis, during model fitting, we encountered a singularity issue, indicating perfect multicollinearity. To address this issue, we removed the intercept for electrodes from the MMN model. We chose to remove the intercept for electrodes rather than the intercept for participant based on model fit diagnostics (model with participant as a random effect: AIC = 3436.31; model with electrode as a random effect: AIC = 3610.86). Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality in either of the two analyzed time window data sets. *P*-values were obtained by likelihood ratio tests of the full model with the main effect of language against the model with no main effects.

Additionally, in order to compare the results across two different learning settings, we conducted a statistical analysis with data pooled from the study of Kędzierska et al. (2023) (naturalistic learners) and from the current study (formal learners). Given that in the study of Kędzierska et al. (2023), 32 electrodes were used,[4] we selected only these electrodes from the fronto-central region which were used in both experiments (C3, C4, Cz, F3, F4, Fz, FC1, FC2, FCz). We constructed a model with Setting (i.e., naturalistic and formal) and Language (i.e., Polish, English and Norwegian) set as fixed effects and with the intercepts for participants and electrodes included as random effects. The same analysis was conducted twice: in the MMN time window (100–200 ms) and in the LDN time window (350–800 ms). *P*-values were obtained by likelihood ratio tests of the full model with the setting × language interaction effect against the model with two main effects.

4. Results

4.1. Behavioral tests results

On average, the participants answered correctly to 73.64% (range: 50.00%–100.00%, *SD* = 16.20%) of the open questions concerning the content of the cartoon they watched. This suggests that during the task, they focused on the movie.

As far as the gating task is concerned, following Kędzierska et al. (2023), we determined both the participants' overall accuracy (range: 56.25%–100.00%, *M* = 79.83%, *SD* = 12.34%) as well as the mean 'gate' at which the word was correctly recognized and the decision was not changed afterward (range: 6.7–10, *M* = 8.53, *SD* = 0.92).

4.2. ERP results

As expected, the deviant sounds in all three languages elicited the MMN-LDN effect. The grand average ERPs and mean voltage difference maps are shown in Figure 1 and Figure 2, respectively. Descriptive statistics for the analyzed conditions in the two-time windows of interest are presented in Table 3.

MMN

In the 100–200 ms time window, we calculated the difference wave between deviant and standard conditions for each participant and each electrode included in the analysis. We then conducted a linear mixed effects analysis of the relationship between the languages and the deviant minus standard difference, reflecting the size of the MMN effect (recall Table 3 for the descriptive statistics for each language condition). Model comparison conducted on this data set revealed a statistically significant main effect of language (χ^2 (2) = 122.06; $p < .001$). Tukey-based pairwise comparisons (see Table 4) revealed that the differences in the MMN effect were statistically significant for each language pair, i.e., English–Norwegian (Estimate = −0.370, $p < .01$), English–Polish (Estimate = 0.783, $p < .001$) and Norwegian–Polish (Estimate = 1.153, $p < .001$) (see Figure 3 and Figure 4). The results of all statistically significant pairwise comparisons for language contrasts are presented in Table 4.

LDN

In the 350–800 ms time window selected to measure the LDN effect, we also conducted a linear mixed effects analysis of the relationship between the languages and the deviant minus standard difference, reflecting the size of the LDN effect (recall Table 3 for the descriptive statistics for each language condition). Model comparison conducted on this data set revealed a statistically significant main effect of language (χ^2 (2) = 26.57; $p < .001$). Tukey-based pairwise comparisons (recall Table 4) revealed that the deviant minus standard difference was statistically significant for the Polish–English pair (Estimate = 0.60, $p < .001$) and the Polish–Norwegian pair (Estimate = 0.60, $p < .001$), but not for the English– Norwegian pair (Estimate = 0.00, $p = .99$) (recall Figure 4).

4.3. Correlation tests

In order to test whether the size of the observed MMN and LDN effects correlated with the participants' AoA, proficiency or phonological aptitude, we conducted six additional linear mixed-effect analyzes, each evaluating the impact of one of the three predictor variables mentioned above on one of the two investigated ERP effects. We report the results of these analyses in the sections below and present the findings visually in Figure 5.

Age of acquisition

To determine if the AoA was related to the MMN or the LDN effect, we conducted a linear mixed effect analysis with the self-reported

[3]In order to distinguish the MMN response from the N1 component, we conducted an additional analysis focused on the comparison of MMN/MMP responses recorded at fronto-central sites and those recorded at the mastoid sites. The results of this analysis are reported in the Supplementary Material (Appendix S2).

[4]While Kędzierska et al. (2023) used a different online reference electrode (FCz), we used the same offline re-reference procedure (i.e., the average of right and left mastoid bones, approximated from TP7 and TP8).

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/S136672892400066X

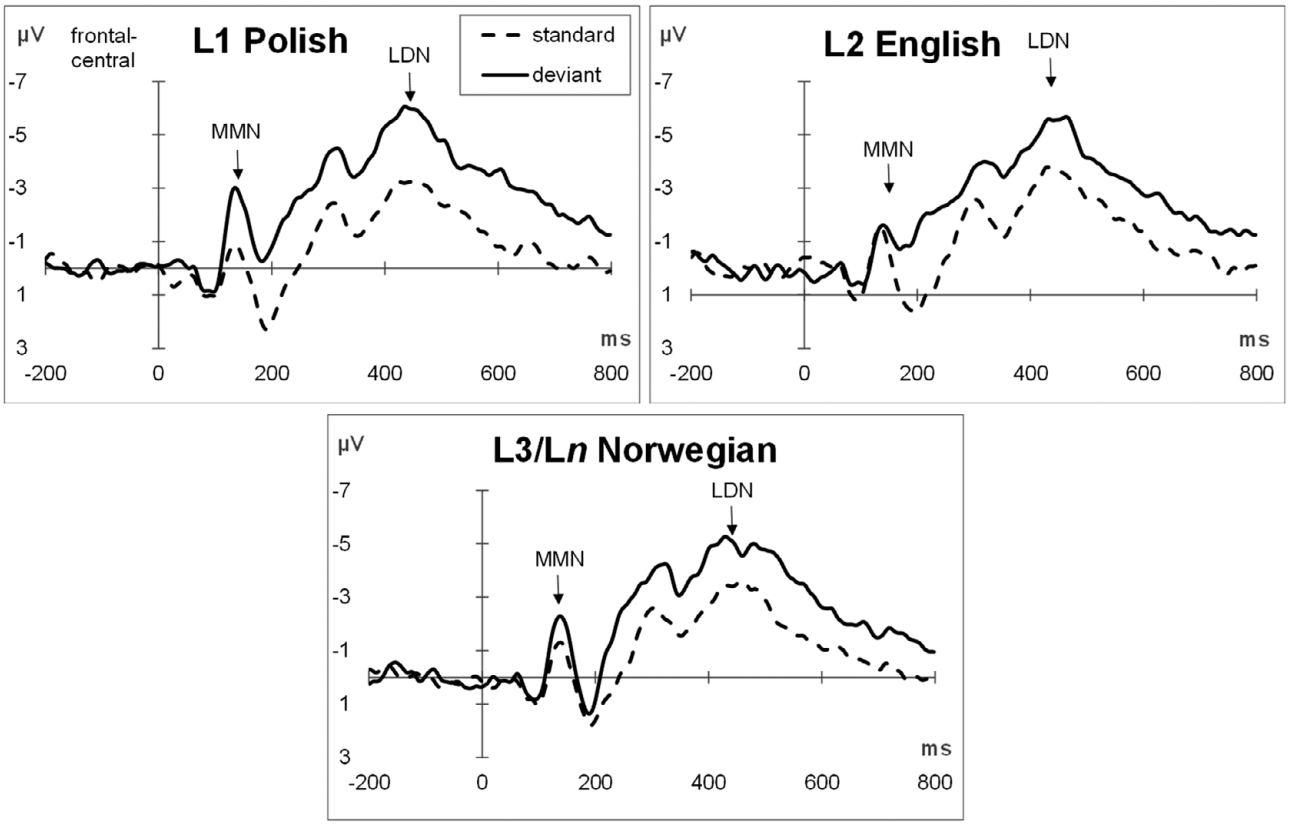

Figure 1. The grand average ERPs time-locked to the onset of the standard (dashed line) and deviant (solid line) for Polish (A), English (B) and Norwegian (C). The figures present the averaged data obtained from 15 EEG electrodes included in the analyses (i.e., F1, F2, F3, F4, Fz, FC1, FC2, FC3, FC4, FCz, C1, C2, C3, C4 and Cz).

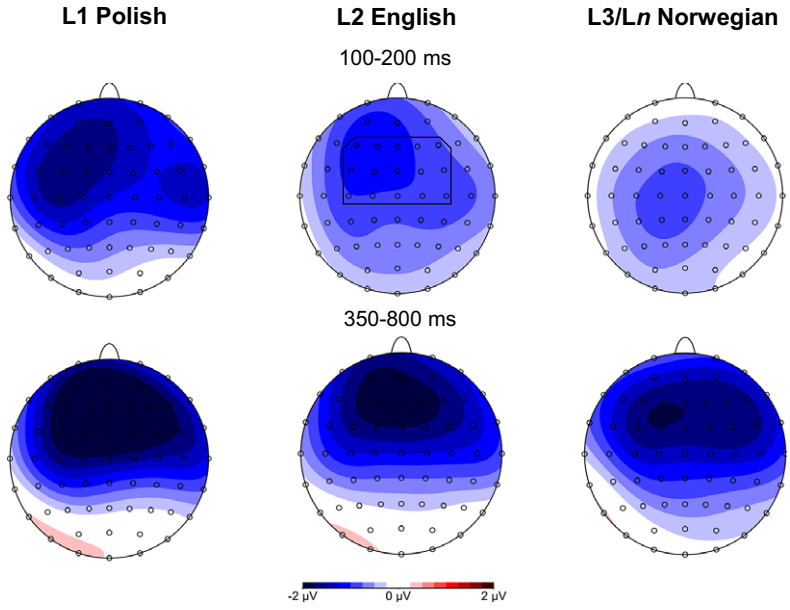

Figure 2. Mean voltage difference maps (deviant minus standard) for Polish (left), English (middle) and Norwegian (right) in the 100–200 ms (upper panel) and 350–800 ms (bottom panel) time windows. The upper middle map presents the electrodes included in the analyses (i.e., F1, F2, F3, F4, Fz, FC1, FC2, FC3, FC4, FCz, C1, C2, C3, C4 and Cz).

ages of acquisition for English and Norwegian as a predictor variable and the average deviant minus standard amplitude value obtained for each participant in the respective language conditions as a criterion variable. An intercept for participants was included as a random effect in all the models reported in this section. We found out that the AoA did not predict the size of the MMN effect (χ^2 (1) = 2.22; p = .14, R^2 = .05) nor of the LDN effect (χ^2 (1) = 0.17; p = .68, R^2 = .00).

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/S136672892400066X

Table 3. Descriptive statistics for the experimental conditions: standard/deviant and Polish/English/Norwegian and for the MMN effect expressed in terms of the deviant minus standard difference in the three language conditions

	emmean	*SE*	*df*	lower.CL	upper.CL
100–200 ms					
L1 Polish					
Standard	0.54	.37	24.2	−0.23	1.30
Deviant	−1.27	.37	24.2	−2.04	−0.51
Deviant–standard	−1.81	.17	27.8	−2.15	−0.31
L2 English					
Standard	0.18	.37	24.2	−0.58	0.95
Deviant	−0.85	.37	24.2	−1.61	−0.08
Deviant–standard	−1.03	.17	27.8	−1.36	−0.68
L3/L*n* Norwegian					
Standard	0.20	.37	24.2	−0.56	0.97
Deviant	−0.46	.37	24.2	−1.22	0.31
Deviant–standard	−0.66	.17	27.8	−0.99	−0.31
350–800 ms					
L1 Polish					
Standard	−1.42	.29	27.5	−2.01	−0.82
Deviant	−3.61	.29	27.5	−4.20	−3.02
Deviant–standard	−2.19	.27	24.7	−2.75	−1.63
L2 English					
Standard	−1.59	.29	27.5	−2.19	−1.00
Deviant	−3.19	.29	27.5	−3.78	−2.59
Deviant–standard	−1.59	.27	24.7	−2.16	−1.03
L3/L*n* Norwegian					
Standard	−1.56	.29	27.5	−2.15	−0.96
Deviant	−3.15	.29	27.5	−3.75	−2.56
Deviant–standard	−1.60	.27	24.8	−2.16	−1.04

Phonological aptitude

Further, we checked whether the size of the observed ERP effects correlated with the participants' phonological aptitude operationalized in terms of the result of the gating task. We conducted a linear mixed effect analysis with the participants' overall accuracy (M = 79.83%, SD = 12.34%) as a predictor variable, but no statistically significant results were observed for the MMN effect (χ^2 (1) = 0.29; p = .60, R^2 = .00) nor for the LDN effect (χ^2 (1) = 2.07; p = .15, R^2 = .11). In addition, we used the mean 'gate' at which the word was correctly recognized (M = 8.53, SD = 0.92) as a predictor variable, but this analysis yielded no statistically significant results (MMN: χ^2 (1) = 0.61; p = .43, R^2 = .01; LDN: χ^2 (1) = 2.17; p = .14, R^2 = .11).

Table 4. Pairwise comparisons for the MMN and LDN effects expressed in terms of the deviant minus standard difference in the three language conditions: Polish, English and Norwegian. Degrees-of-freedom method: Kenward-Roger; p-value adjustment: Tukey method for comparing a family of 3 estimates

Compared Conditions	Estimate	*SE*	*df*	t.ratio	*p*-value
100–200 ms					
Contrast: English–Polish	0.78	0.103	965	7.581	<.001
Contrast: Norwegian–Polish	1.15	0.103	965	11.157	<.001
Contrast: English–Norwegian	−0.37	0.103	965	−3.581	.001
350–800 ms					
Contrast: English–Polish	0.60	0.133	965	4.498	<.001
Contrast: Norwegian–Polish	0.59	0.133	965	4.462	<.001
Contrast: English–Norwegian	0.00	0.133	965	0.033	.99

Proficiency

In the final analysis, we used the participants' scores obtained in English and Norwegian proficiency tests as a predictor variable determining their ERP responses in the English and Norwegian language conditions. We found out that proficiency influenced the MMN effect (χ^2 (1) = 3.68; p = .05, R^2 = .08), but not the LDN effect (χ^2 (1) = 1.72; p = .19, R^2 = .22).

4.4. Settings comparison

The results of the comparative analysis revealed a statistically significant Setting × Language interaction effect in the MMN time window (χ^2 (2) = 6.15; p < .05). Tukey-based pairwise comparisons revealed that the MMN effect was significantly different for each language pair in the case of formal learners, i.e., English–Norwegian (Estimate = −0.402, p < .01), English–Polish (Estimate = 0.779, p < .001) and Norwegian–Polish (Estimate = 1.181, p < .001), and for two language pairs, i.e., English–Norwegian (Estimate = −0.440,

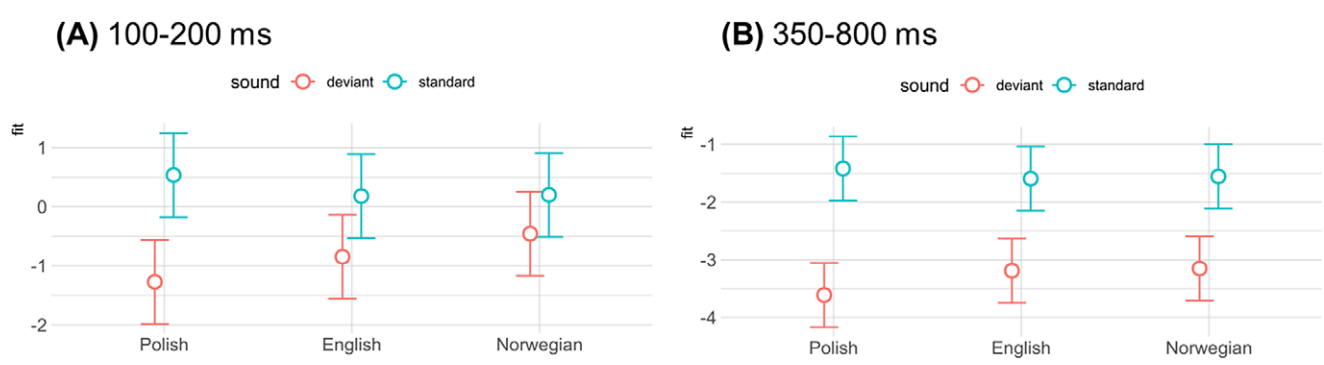

Figure 3. Mean amplitude values (in μV) observed in each standard/deviant condition and each target language within the 100–200 ms time window (panel A) and the 350–800 ms time window (panel B). Bars indicate standard deviations.

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/S136672892400066X

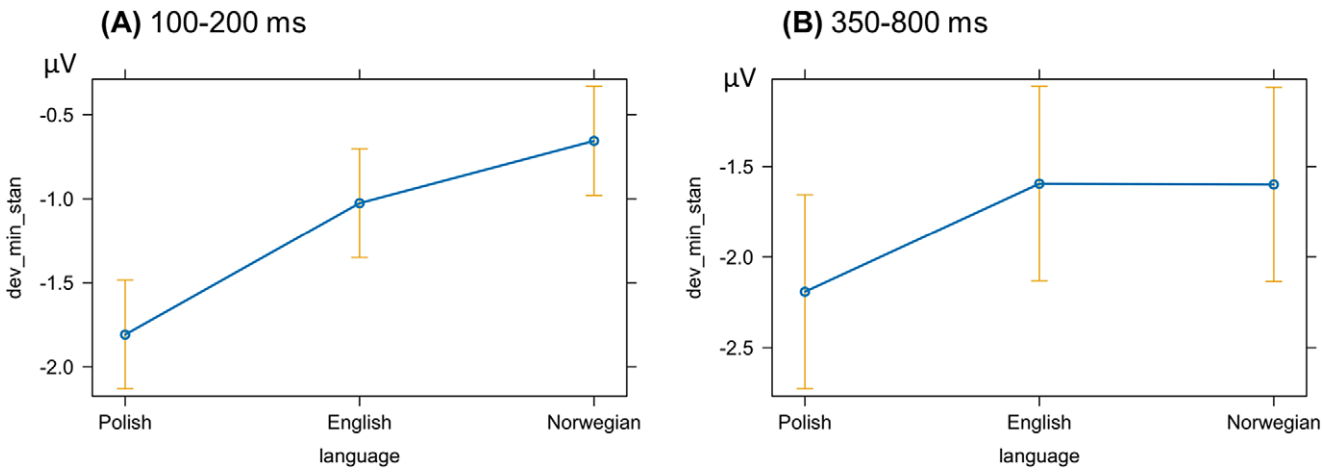

Figure 4. Mean amplitude differences, i.e., deviant minus standard (in μV) for each target language in the 100–200 ms time window (panel A), the 350–800 ms time window (panel B). Bars indicate standard deviations.

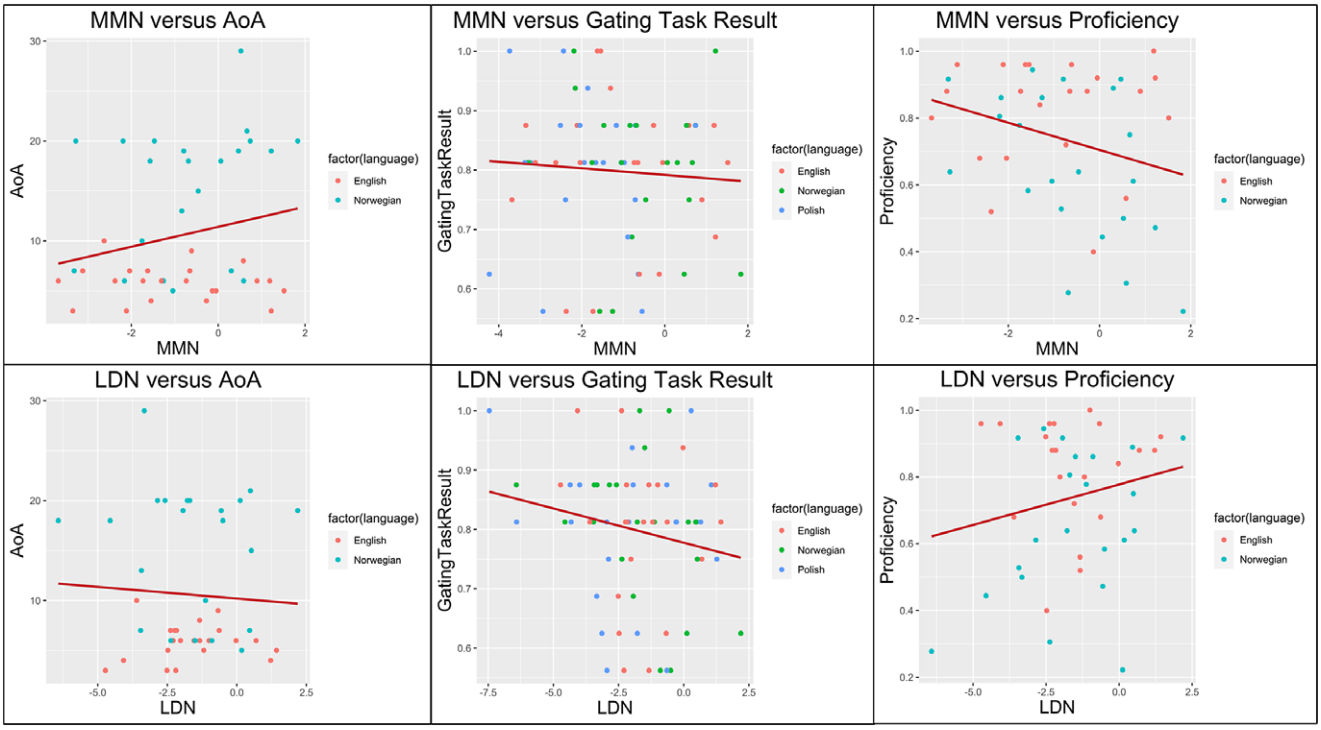

Figure 5. The results of correlational tests assessing whether AoA, phonological aptitude and proficiency predicted the size of MMN and LDN effects.

$p < .01$) and Norwegian–Polish (Estimate = 0.775, $p < .001$) in the case of naturalistic learners. The results of all pairwise comparisons for language and setting contrasts are presented in Table 5.

In the LDN time window, we only observed a statistically significant main effect of Language (χ^2 (2) = 33.31; $p < .001$). Tukey-based pairwise comparisons revealed that Polish stimuli elicited a significantly larger LDN effect when compared with English and Norwegian stimuli (English–Polish: Estimate = 0.6028, $p < .001$; Norwegian–Polish: Estimate = 0.5874, $p < .001$). We did not find any statistically significant difference between the two non-native languages, though (English–Norwegian: Estimate = 0.0154, $p = .99$). Figures presenting the comparison of ERP effects obtained for formal learners (the current study) and naturalistic learners (the study of Kędzierska et al., 2023) are included in the Supplementary Material (Appendix S3).

4.5 Comparison of differences in phonetic parameters and MMN for language pairs

In order to relate differences in phonetic parameters to differences in MMN obtained for language pairs, we can compare data in Table S2 (Appendix S1), which summarizes formant differences for language pairs to the results in Table 4, which summarizes pairwise comparisons for the MMN and LDN effects expressed in terms of the deviant minus standard difference in the three language conditions. For the time window between 100 and 200 ms,

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/S136672892400066X

Table 5. Pairwise comparisons for the MMN effect (100–200 ms time window) expressed in terms of the deviant minus standard difference in the three language conditions: Polish, English and Norwegian and two learnings settings: formal (the current study) and naturalistic (Kędzierska et al., 2023). Degrees-of-freedom method: Kenward-Roger; p-value adjustment: Tukey method for comparing a family of 3 estimates

Compared Conditions	Estimate	*SE*	*df*	*t*.ratio	p-value
Naturalistic setting					
Contrast: English–Polish	0.34	0.143	1088	2.342	.051
Contrast: Norwegian–Polish	0.76	0.143	1088	5.406	<.001
Contrast: English–Norwegian	−0.44	0.143	1088	−3.064	< .01
Formal setting					
Contrast: English–Polish	0.78	0.137	1088	5.697	<.001
Contrast: Norwegian–Polish	1.18	0.137	1088	8.636	<.001
Contrast: English–Norwegian	−0.40	0.137	1088	−2.939	<.01

the largest difference was observed for Norwegian–Polish, then English–Polish and lastly English–Norwegian. Out of all the phonetic parameters that were taken into account, i.e., F1 (related to differences in vowel height), F2 (related to the tongue advancement) and F3 or F3–F2 (related to lip rounding) and Euclidean distances between vowels (i.e., Euclidean distances calculated for either F1 and F2 or F1, F2 and F3), it turns out that MMN differences were most closely related to differences in F1. The hierarchy of differences was not organized in the same order for other phonetic parameters.

5. Discussion

In the globalized world, it is possible to learn a new language through various means and modalities. In the absence of opportunity to analyze every possible manner of learning a new language, we decided to narrow down our research questions so that they applied to two radically different groups of language learners, i.e., those acquiring a foreign language in a formal classroom setting and those learning it abroad in a naturalistic setting. We wondered whether significant differences would emerge in terms of L1 Polish, L2 English, and L3/L*n* Norwegian phoneme processing among formal L2 English and L3/L*n* Norwegian learners, and whether we would observe any differences between this group of learners, when compared with an independently recorded group of naturalistic L3/L*n* Norwegian learners (Kędzierska et al., 2023). To achieve this goal, we used a passive oddball paradigm, where the vowels in each investigated language were the same as the stimuli used by Kędzierska et al. (2023) and the procedure was almost identical to those employed in the study involving naturalistic Norwegian learners. The participants in the current experiment acquired both foreign languages under investigation in a formal classroom setting. For this group, the difference between L2 English and L3/L*n* Norwegian learning history mostly lies in the much earlier AoA for L2 English (i.e., 5.86 years on average), when compared with L3/L*n* Norwegian (i.e., at 20.27 years on average). Unlike the naturalistic learners in the study of Kędzierska et al. (2023), in the case of the current participants, the L3/L*n* acquisition proceeded in a very intensive way during academic-level courses. What seems vital in the comparison of these two groups of learners is that the naturalistic learners acquired both foreign languages later than the instructed group (i.e., at around 9.48 years and 27.33 years, respectively). What is more, the naturalistic group predominantly reported to use English much more frequently than Norwegian in their everyday lives (Kędzierska et al., 2023, p. 9).

The study aimed to answer several research questions, with the most general one concerning the comparison of the MMN amplitudes in the native language (i.e., Polish) with respect to the non-native languages (i.e., L2 English and L3/L*n* Norwegian). Previous studies in bilingual phoneme processing have shown that the MMN effect is typically smaller for non-native when compared with native phoneme contrasts, especially if the investigated languages are acquired in a formal setting (e.g., Jakoby et al., 2011; Liang & Chen, 2022; Song & Iverson, 2018). Therefore, we hypothesized that the MMN effect will be significantly larger in L1 Polish when compared with L2 English and L3/L*n* Norwegian. This prediction was confirmed in the instructed learning context in the present study as the amplitude of the MMN was the greatest in L1 Polish, smaller in L2 English and the smallest in L3/L*n* Norwegian, with all pairwise comparisons yielding statistically significant results. However, in the previous study investigating phonemic perception in the naturalistic context for language learners recorded in Norway, Kędzierska et al. (2023) observed no statistically significant difference between L1 Polish and L2 English. Our analysis of data pooled from the two studies further confirmed this finding. This result shows that despite higher AoA, the naturalistic group seems to have developed more native-like mechanisms of phoneme processing, when compared with the instructed group. Quite interestingly (and perhaps somewhat counterintuitively), the difference is only visible in L2 English, though. Since the naturalistic participants were more dominant in L2 English than in L3/L*n* Norwegian (i.e., they predominantly used English for communication, especially in the work environment, despite living in Norway), this difference, in our view, should be attributed to the influence of immersive learning context on the ability to discriminate non-native phonemic contrasts. This conclusion is also in accordance with the results obtained by Winkler et al. (1999), who showed that naturalistic learners of Finnish exhibited a native-like MMN response to Finnish vowel contrasts.

In terms of the differences between L2 English and L3/L*n* Norwegian, a weaker MMN effect in the L3/L*n* also signals the importance of – very broadly understood – language dominance as a factor greatly contributing to effective phoneme discrimination. Early age of acquisition and higher language proficiency seem to considerably facilitate this task, as evidenced by the results of several correlation analyses in the current study (recall Section 3.3) as well as those conducted previously by Kędzierska et al. (2023). As in the case of numerous previous studies on non-native phoneme discrimination, in the current study, the MMN was followed by the LDN effect, which is – among other suggested explanations – believed to indicate whether memory traces associated with a specific phonemic representation have been formed successfully (Barry et al., 2009; Jakoby et al., 2011). For this response, we also observed statistically significant differences between L1 Polish and the two non-native languages in question in the instructed learning context. In contrast to the MMN, we found no statistically significant difference between L2 English and L3/L*n* Norwegian, though. In this way, we have replicated the results of Kędzierska et al. (2023), who also observed no such difference in naturalistic learners – a finding further corroborated by the results of a comparative statistical analysis with data pooled from the two studies. This result supports an idea still debated in the literature (e.g., Deroche et al.,

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/S136672892400066X

2023; Halliday et al., 2014) that the two ERP components – the MMN and the LDN – indicate two distinct processes.

In light of the current results, while the MMN might (in consistence with previous findings) be assumed to index auditory discrimination, the functional significance of the LDN should be viewed as going beyond the mere continuation of the earlier effect (i.e., the so-called 'late MMN' interpretation), since the two components do not show a similar pattern. The current results also seem to speak against the interpretation of the LDN as a response evoked by the re-orientation of attention after being distracted by a deviant sound (Shestakova et al., 2003; Wetzel et al., 2006). Rather, the LDN might be considered an index of successful formation of phonological representations (Barry et al., 2009), which – at least in the current study – remained unaffected by the foreign language status (L2 and L3/L*n*) or learning context (instructed and naturalistic). This would suggest that the processing differences between the two non-native languages are better reflected at the pre-attentional level in the MMN time window, and – assuming a sufficient amount of training has been received – might further shade at a higher order cognitive level.

A more detailed examination of the acoustic parameters of the vowel stimuli was conducted to ensure that the observed effects are not solely attributable to acoustic differences between standards and deviants in each language block. Since we are dealing with natural spoken languages, it is impossible to designate equidistant stimuli. When Euclidean distances in Bark between stimuli in each language block were examined, the distance was the largest in English (4.77), then in Polish (2.31) and the smallest in Norwegian (1.51). The MMN effects in the 100–200 ms time window, as evaluated on the basis of estimates in pairwise comparisons between experimental conditions, were, however, the largest for Polish (−1.81), then for English (−1.03) and the smallest for Norwegian (−0.66). This means that although the Euclidean distance between Polish stimuli was nearly twice as small as between English stimuli, it was the native phonemic contrast that evoked larger MMN effects. In similar studies (Kędzierska et al., submitted; Kędzierska et al., 2023), Euclidean distances did not determine the MMN effects, either. In Kędzierska et al. (2023), a study with the same stimuli and a different participant group of naturalistic language learners, the MMN effects were similar for L1 Polish and L2 English, but smaller for L3 Norwegian. In Kędzierska et al. (submitted), the Euclidean distances between stimuli in all languages were fairly balanced, and we obtained similar results when it comes to the hierarchy of MMN effects, suggesting that these were not pure acoustic differences between the stimuli in a given language block that influenced the MMN response.

Also, although the stimuli were presented to the participants in separate language blocks, in the statistical analysis, we compared the MMN effect size across languages. When we attempted to compare distances between vowel stimuli in Bark across languages, it turned out that out of all phonetic parameters taken into account (i.e., F1, F2, F3, F3–F2 and Euclidean distances based on F1 and F2 or F1, F2 and F3), it was the distance in F1 (see Appendix S1) that best reflected the hierarchy of MMN differences (see Table S2 in Appendix S1): both the distance and the MMN effect were largest for the Norwegian minus Polish condition, followed by English minus Polish and English minus Norwegian. F1 is related to vowel height, and vowel height distinctions in world languages are equal to or more numerous than backness distinctions (Crothers, 1978). The former turned out to be relatively easily perceivable (Balas, 2018). The lack of a straightforward relationship between other phonetic parameters, especially Euclidean distances and MMN effects, suggests that the MMN effects that we observed were not artifacts of acoustic effects, but were rather attributable to language status (native versus non-native) and AoA. So far, we have shown that the largest Euclidean distance between L2 vowels did not entail the biggest MMN effects. Future research should examine other combinations of languages and vowel contrasts, especially with the largest Euclidean distance in L3/L*n* when compared with L2 and L1. In such studies, other vowel pairs could be examined in order to disentangle the effects of language status (operationalized as L2 or L3/L*n*) and the effects associated with the characteristics of the presented auditory stimuli. However, it is not viable to include all combinations of distances between stimuli in a single study due to inherent limitations of natural languages.

One limitation of the current study is that it involves a relatively small sample size stemming from the overall low population of Polish–English–Norwegian trilinguals. This also resulted in a specific acquisition order with L2 English and L3 Norwegian. The reverse combination would hardly be feasible to find among adult formal learners due to the well-established status of English as the most frequent L2 choice. Moreover, assessing speech-specific capabilities could be extended to other tasks, e.g., a word identification or a lexical decision task, rather than only the gating task, which should also ideally involve phoneme pairs in all three languages under investigation. The reason why we decided to conduct the task only in one language was motivated by the overall long duration of the study. We focused on English since we predicted the Polish vowel contrasts might be too easy to detect for native Polish speakers, given the generally modest repertoire of Polish vowels, while English was chronologically the first and more advanced foreign language of our participants.

6. Conclusions

The main objective of this study was to investigate the processing of vowel contrasts by multilingual learners, a topic particularly relevant in times of rapid globalization, the strengthening status of English as a *lingua franca*, and numerous other factors that favor increasing linguistic heterogeneity. The novelty of our investigation includes the selection of trilingual participants who learnt both foreign languages in a formal setting and the analysis of the results in the light of those obtained earlier for a group of naturalistic/mixed learners of the same languages. Altogether, our research supports the idea that foreign language status (i.e., L2 vs. L3/L*n*) influences early auditory processing. However, such factors as learning context and language proficiency can undeniably affect the efficiency of phoneme discrimination in a non-native language. Therefore, under certain circumstances, the contrasts between native and non-native languages may considerably diminish, possibly leading to native-like phoneme discrimination in the most advanced learners and/or for the most dominant language.

The novelty of the current research mostly relies on the inclusion of various learners' groups: all studies to date have investigated non-native phoneme perception from the perspective of one acquisition setting – be it either naturalistic or instructed. The current study is, to the best of our knowledge, the first ERP study which focused on a purely formal (i.e., classroom-instruction) learners of two non-native languages. We then analyzed the results in the light of the ones obtained earlier for exactly the same experimental material, but processed by naturalistic/mixed learners (Kędzierska et al., 2023). An advantage observed for the naturalistic group who seemed to process the L1 and L2 sounds

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms.
https://doi.org/10.1017/S136672892400066X

equally effectively yet again shows that non-native phonemic processing benefits from frequent use in natural communicative contexts. However, it should not so much lead to a detrimental division into 'better' and 'worse' acquisition settings, but rather should inform educational decisions of those whom they concern. Undoubtedly, phonemic processing and, consequently, successful phonemic production, is one of the most relevant skills in which any learner strives to achieve competence. Understanding the determinants that enhance the efficiency of non-native phonemes acquisition ought to exert more impact on the pedagogical decisions undertaken by educators, as well as on the choices made by learners or their parents.

Supplementary material. To view supplementary material for this article, please visit http://doi.org/10.1017/S136672892400066X.

Data availability statement. The data that support the findings of this study are openly available in the OSF repository at the following link: https://osf.io/2956a/.

Acknowledgements. We thank Zuzanna Cal and Tristan Czarnecki-Verner for their assistance with the EEG recordings and Kamil Kaźmierski for his comments on data analysis.

Funding statement. The research leading to these results has received funding from the Norwegian Financial Mechanism 2014–2021 project number 2019/34/H/HS2/00495.

Competing interest. The authors declare none.

References

Balas, A. (2018). *Non-native vowel perception: The interplay of categories and features*. Wydawnictwo Naukowe UAM.

Barry, J. G., Hardiman, M. J., & Bishop, D. V. (2009). Mismatch response to polysyllabic nonwords: A neurophysiological signature of language learning capacity. *PLoS One*, **4**, e6270. https://doi.org/10.1371/journal.pone.0006270

Bates, D. M., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes.* R package version 0.999999-0.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, **109** (2), 775–794. https://doi.org/10.1121/1.1332378

Best, C. T., & Tyler, M. D. (2007) Non-native and second language speech perception: Commonalities and complementarities. In M. J. Munro & O. S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13–34.) John Benjamins.

Bjelaković, A. (2016). The vowels of contemporary RP: vowel formant measurements for BBC newsreaders. *English Language and Linguistics*, **21**(3), 501–532. https://doi.org/10.1017/s1360674316000253

Bjelaković, A. (2017). The vowels of contemporary RP: Vowel formant measurements for BBC newsreaders. *English Language and Linguistics*, **21**(3), 501–532. https://doi.org/10.1017/S1360674316000253

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, **5**(9/10), 341–345.

Cabrelli Amaro, J. (2017). Testing the phonological permeability hypothesis: L3 phonological effects on L1 versus L2 systems. *International Journal of Bilingualism*, **21**(6), 698–717. https://doi.org/10.1177/1367006916637287

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

Crothers, J. (1978). Typology and universals of vowel systems. In J. Greenberg, C. Ferguson, & E. Movarvcsik (Eds.), *Universals of human language* (Vol. **2**). Stanford University Press.

Deroche, M. L. D., Wolfe, J., Neumann, S., Manning, J., Towler, W., Alemi, R., Bien, A. G., Koirala, N., Hanna, L., Henry, L., & Gracco, V. L. (2023). Auditory evoked response to an oddball paradigm in children wearing cochlear implants. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, **149**, 133–145. https://doi.org/10.1016/j.clinph.2023.02.179

Díaz, B., Mitterer, H., Broersma, M., Escera, C., & Sebastián-Gallés, N. (2016). Variability in L2 phonemic learning originates from speech-specific capabilities: An MMN study on late bilinguals. *Bilingualism: Language and Cognition*, **19**(5), 955–970. https://doi.org/10.1017/S1366728915000450

Di Dona, G., Scaltritti, M., & Sulpizio, S. (2022). Formant-invariant voice and pitch representations are pre-attentively formed from constantly varying speech and non-speech stimuli. *European Journal of Neuroscience*, **56**, 4086–4106. https://doi.org/10.1111/ejn.15730

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. LOT.

Escudero, P. (2009). The linguistic perception of similar L2 sounds. In P. Boersma & S. Hamann (Eds.), *Phonolgoy in perception* (pp. 152–190). Walter de Gruyter.

Escudero, P. & Boersma, P.(2004). Bridging the gap between L2 speech perception and phonological theory. *Studies in Second Language Acquisition*, **26**, 551–585. https://doi.org/10.1017/S0272263104040021

Flege, J.E. (1995). Second language speech learning: Theory, problems, findings. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.

Flege, J. E., & Bohn, O. S. (2021). The revised Speech Learning Model. In R. Wayland (Ed.), *Second language speech learning. Theoretical and empirical progress* (pp. 84–118). Cambridge University Press.

Flege, J. E., Munro, M. J. & Fox, R. A. (1994). Auditory and categorial effects on cross-language vowel perception. *Journal of the Acoustical Society of America* **95**: 3623–3641. https://doi.org/10.1121/1.409931

Halliday, L. F., Barry, J. G., Hardiman, M. J., & Bishop, D. V. (2014). Late, not early mismatch responses to changes in frequency are reduced or deviant in children with dyslexia: an event-related potential study. *Journal of Neurodevelopmental Disorders*, **6**(1), 21. https://doi.org/10.1186/1866-1955-6-21

Jakoby, H., Goldstein, A., & Faust, M. (2011). Electrophysiological correlates of speech perception mechanisms and individual differences in second language attainment. *Psychophysiology*, **48**, 1516–1530. https://doi.org/10.1111/j.1469-8986.2011.01227.x

Jassem, W. (2003). Polish. *Journal of the International Phonetic Association*, **33**, 103–107. https://doi.org/10.1017/S0025100303001191

Kaan, E. (2007). Event related potentials and language processing: A brief overview. *Language and Linguistic Compass*, **1**(6), 571–591. https://doi.org/10.1111/j.1749-818x.2007.00037.x

Kędzierska, H., Rataj, K., Balas, A., Cal, Z., Castle, C., & Wrembel, M. (2023). Vowel perception in multilingual speakers: ERP evidence from Polish, English and Norwegian. *Frontiers in Psychology*, **14**(1270743). https://doi.org/10.3389/fpsyg.2023.1270743

Kędzierska, H., Rataj, K., Balas, A., & Wrembel, M. (submitted). MMN asymmetries in trilingual listeners.

Kopečková, R. (2015). Differences in the perception of English vowel sounds by child L2 and L3 learners. In U. Gut, R. Fuchs & E. Wunder (Eds.), *Universal or diverse paths to English phonology* (pp. 71–90). De Gruyter Mouton. https://doi.org/10.1515/9783110346084-005

Kristoffersen, G. (2000). *The phonology of Norwegian*. Oxford University Press.

Kujala, A., & Näätänen, R. (2003). Auditory environment and change detection as indexed by the mismatch negativity (MMN). In J. Polich (Ed.), *Detection of change* (pp. 1–22). Springer.

Levy, E. S. & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, **36**, 141–157. https://doi.org/10.1016/j.wocn.2007.03.001

Liang, L., & Chen, B. (2022). The non-native phonetic perception mechanism utilized by bilinguals with different L2 proficiency levels. *International Journal of Bilingualism*, **26**(3), 368–386. https://doi.org/10.1177/13670069211058275

Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language history questionnaire (LHQ3): an enhanced 709 tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, **23**, 938–944.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries.

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S136672892400066X

Journal of Experimental Psychology, **54**(5), 358–368. https://doi.org/10.1037/h0044417

Luck, S. J., & **Gaspelin, N.** (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, **54**, 146–157. https://doi.org/10.1111/psyp.12639

Lloyd-Smith, A. (2023). Perceived foreign accent in L3 English: The effects of heritage language use. *International Journal of Multilingualism*, **20**(3), 753–767. https://doi.org/10.1080/14790718.2021.1957899

Maddieson, I. (2013). Front rounded vowels. In M. S. Dryer & M. Haspelmath (Eds.) *WALS Online (v2020.3)*. https://doi.org/10.5281/zenodo.7385533; https://wals.info/chapter/11

Muñoz, C. (2008). Symmetries and Asymmetries of Age Effects in Naturalistic and Instructed L2 Learning, *Applied Linguistics*, **29**(4), 578–596.

Näätänen, R., **Lehtokoski, A.**, **Lennes, M.**, **Cheour, M.**, **Huotilainen, M.**, **Iivonen, A.**, **Vainio, M.**, **Alku, P.**, **Ilmoniemi, R. J.**, **Luuk, A.**, **Allik, J.**, **Sinkkonen, J.**, & **Alho, K.** (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, **385**(6615), 432–434. https://doi.org/10.1038/385432a0

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

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, **9**, 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Peirce, J., **Gray, J. R.**, **Simpson, S.**, **MacAskill, M.**, **Höchenberger, R.**, **Sogo, H.**, **Kastman, E.**, & **Lindeløv, J. K.** (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, **51**(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

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

Peltola, M. S., **Tamminen, H.**, **Toivonen, H.**, **Kujala, T.**, & **Näätänen, R.** (2012). Different kinds of bilinguals – different kinds of brains: the neural organisation of two languages in one brain. *Brain and Language*, **121**(3), 261–266. https://doi.org/10.1016/j.bandl.2012.03.007

Puig-Mayenco, E., **González Alonso, J.**, & **Rothman, J.** (2020). A systematic review of transfer studies in third language acquisition. *Second Language Research*, **36**(1), 31–64. https://doi.org/10.1177/0267658318809147

R Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Sebastián-Gallés, N., & **Baus, C.** (2005). On the Relationship Between Perception and Production in L2 Categories. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 279–282). Lawrence Erlbaum Associates Publishers.

Sebastián-Gallés, N., & **Soto-Faraco, S.** (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, **72**(2), 111–123. https://doi.org/10.1016/S0010-0277(99)00024-4

Shestakova, A., **Huotilainen, M.**, **Ceponiene, R.**, & **Cheour, M.** (2003). Event-related potentials associated with second language learning in children. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*, **114**(8), 1507–1512. https://doi.org/10.1016/s1388-2457(03)00134-2

Song, J., & **Iverson, P.** (2018). Listening effort during speech perception enhances auditory and lexical processing for non-native listeners and accents. *Cognition*, **179**, 163–170. https://doi.org/10.1016/j.cognition.2018.06.001

Tyler, M. D., **Best, C.T.**, **Faber, A.** & **Levitt, A.G.** 2014. Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, **71**, 4–21. https://doi.org/10.1159/000356237

Upton, C., **Kretzchmar, W.** & **Konopka, R.** (2003). *Oxford Dictionary of pronunciation of current English*. Oxford: Oxford University Press.

Weckwerth, J., & **Balas, A.** (2019). Selected Aspects of Polish Vowel Formants. In *Approaches to the Study of Sound Structure and Speech* (p. 338–348). Routledge. https://doi.org/10.4324/9780429321757-25

Wells, J. C. (1962). *A study of the formants of the pure vowels of British English*. MA dissertation. University College London.

Wetzel, N., **Widmann, A.**, **Berti, S.**, & **Schröger, E.** (2006). The development of involuntary and voluntary attention from childhood to adulthood: A combined behavioral and event-related potential study. *Clinical Neurophysiology*, **117**, 2191–2203. https://doi.org/10.1016/j.clinph.2006.06.717

Winkler, I., **Kujala, T.**, **Tiitinen, H.**, **Sivonen, P.**, **Alku, P.**, **Lehtokoski, A.**, **Czigler, I.**, **Csépe, V.**, **Ilmoniemi, R. J.**, & **Näätänen, R.** (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, **36**, 638–642.

White, E., **Titone, D.**, **Genesee, F.**, & **Steinhauer, K.** (2017). Phonological processing in late second language learners: The effects of proficiency and task. *Bilingualism: Language and Cognition*, **20**(1), 162–183. https://doi.org/10.1017/S1366728915000620

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials (ERPs) in studies of perception and attention. *Attention, Perception and Psychophysics*, **72**(8), 1–29. https://doi.org/10.3758/APP.72.8.2031

Wottawa, J., **Adda-Decker, M.**, & **Isel, F.** (2022). Neurophysiology of non-native sound discrimination: Evidence from German vowels and consonants in successive French–German bilinguals using an MMN oddball paradigm. *Bilingualism: Language and Cognition*, **25**(1), 137–147. https://doi.org/10.1017/S1366728921000468

Wrembel, M., **Marecka, M.**, & **Kopečková, R.** (2019). Extending perceptual assimilation model to L3 phonological acquisition. *International Journal of Multilingualism*, **16**(4), 513–533. https://doi.org/10.1080/14790718.2019.1583233

Wrembel, M., **Marecka, M.**, **Szewczyk, J.**, & **Otwinowska, A.** (2019). The predictors of foreign-accentedness in the home language of Polish–English bilingual children. *Bilingualism: Language and Cognition*, **22**(2), 383–400. https://doi.org/10.1017/S1366728918000044

Wrembel, M. (2023). Exploring the Acquisition of L3 Phonology: Challenges, New Insights, and Future Directions. In J. Cabrelli, A. Chaouch-Orozco, J. González Alonso, S. Pereira Soares, E. Puig-Mayenco, & J. Rothman (Eds.), *The Cambridge handbook of third language acquisition* (Cambridge handbooks in language and linguistics, pp. 115–141). Cambridge University Press. https://doi.org/10.1017/9781108957823.006

Downloaded from https://www.cambridge.org/core. IP address: 118.169.141.33, on 03 Oct 2026 at 12:51:41, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S136672892400066X