Investigation of long- and short-term adaptations of the bilingual language system to different language environments: evidence from the ERPs

Alba Casado 1,2*, Jonas Walther 1,3*, Agata Wolna 1,4, Jakub Szewczyk 1, Antonella Sorace 5 & Zofia Wodniecka 1

1. Psychology of Language and Bilingualism Lab, Institute of Psychology, Jagiellonian University, Krakow, Poland

2. Mind, Brain, and Behavior Research Center (CIMCYC); Department of Experimental Psychology; University of Granada, Granada, Spain

3. Department of Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany

4. McGovern Institute for Brain Research, MIT, Cambridge, MA, 02139, USA

5. School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, United Kingdom

*Joint first authorship

Address for correspondence

Alba Casado, e-mail – albacasado@ugr.es

Mind, Brain, and Behaviour Research Center (CIMCYC), Department of Experimental Psychology, University of Granada. Campus Cartuja. C/ Profesor Clavera s/n, 18011, Granada, Spain

Zofia Wodniecka, e-mail – zofia.wodniecka@uj.edu.pl

Psychology of Language and Bilingualism Lab, Institute of Psychology, Jagiellonian University in Krakow, Ul. Ingardena 6, 30-060

Abstract

Does a long-term stay in a foreign language country affect word retrieval in our native language? And if so, are the effects reversible? The present study explored the neural correlates of single-word production in the native language and their dynamics due to two types of changes in the language environment: long-term immersion in a foreign language (L2) environment and short-term reimmersion in a native language (L1) environment. We tested Polish-English migrants living in the UK (L2 environment) for an average of ten years and Polish-English controls living in Poland (L1 environment). All participants performed an L1 picture-naming task while we recorded their electrophysiological responses. The migrants were tested before and after visiting the L1 environment, while the controls were tested twice in their L1 environment. Our focus was on two event-related components previously associated with the ease of lexical access: P2 and N300. We found no modulations related to N300, but some in the P2 time window, although their distribution was more frontal than previously reported. There was no main effect of the long-term immersion in the L2 environment, suggesting that the effectiveness of producing words in L1 was similar across the two groups. However, the short-term change in the language environment modulated the early positivity in migrants: smaller frontal positivity was reported in response to picture naming after the short reimmersion in the L1 environment than during the L2 immersion. These results indicate that the short-term changes in the language environment induce modulations in the neural response, which may reflect higher proactive control applied in L1 production during L2 immersion and its reduction after short-term L1 immersion.

Keywords: bilingual lexical access; L2 immersion; L1 reimmersion; P2; N300.

1. Introduction

Moving to a foreign country with a language different from the native one usually requires the speaker to use the second language (L2) daily. Being immersed in an environment where L2 is predominant can affect the speakers' language system. While speakers immersed in an L2 environment have plenty of opportunities to hone their L2, they have fewer opportunities to use their native language (L1), and there is a reduced variability of L1 interlocutors. These two factors are likely responsible for the fact that speakers immersed in L2 environments often experience difficulties producing words in L1 (Ammerlaan, 1996; Hulsen, 2000; Linck et al., 2009; Schmid & Jarvis, 2014; Schmid & Keijzer, 2009; Yagmur et al., 1999), especially low-frequency words (Baus et al., 2013; Botezatu et al., 2021). These difficulties using L1 are usually discussed in the literature as an overall decrease in L1 access or availability. Alternative explanations relate to more fundamental changes in the lexicon or general language knowledge due to L2 immersion. Overall, the phenomenon has been called 'language attrition' (Köpke & Schmid, 2004; Schmid, 2009). Still, the decrease in L1 accessibility due to L2 immersion is not experienced in all linguistic domains such as speech fluency or language comprehension in morphosyntactic processing (Bergmann et al., 2015; Gnitiev & Bátyi, 2022; Schmid, 2009). The decrease in L1 accessibility also does not seem to be permanent, as previous studies showed that reimmersion in the native language environment leads to a return of language skills to a level before L2 immersion (Chamorro et al., 2016; Köpke & Genevska-Hanke, 2018; Linck et al., 2009). However, not all studies report a behavioural effect of long-term L2 immersion on L1 access (Beatty-Martínez et al., 2020; Casado, Walther, et al., 2023; Yilmaz & Schmid, 2012). It is therefore of interest to explore whether the absence of behavioural effects can still be accompanied by signs of difficulty on the neural level.

Given that none of the previous studies have employed electrophysiological indices to track changes due to language immersion in language processing at the brain level, our goal was to fill in the gap in the literature. The electrophysiological (EEG) technique is a tool that allows one to explore brain reactions to stimuli with high time resolution. As a result, event-related potentials (ERPs) can provide a measure of the processing of stimuli even when there is no behavioural change (Bice & Kroll, 2015; McLaughlin et al., 2004). Thus, EEG allows us to investigate whether the difficulty in accessing the native language in migrants is reflected by the ERP markers and whether the same markers are also sensitive to short-term reimmersion to the L1. Consequently, in the current study, we examined the electrophysiological response in the picture naming of Polish-English migrants in an L1 (Polish) and an L2 (English) environment. Below, we briefly review the differing results of studies on how L1 lexical access can be affected by the language environment.

1.1. Previous studies exploring long-term bilinguals' L1 lexical access in different language environments

The data on the effect of L2 immersion are somewhat inconclusive. On the one hand, previous studies reported no differences in the speed of lexical access to L1 when L1 access is compared between groups of bilinguals immersed in L1 and bilinguals immersed in L2 for a long time, at least when measured using a relatively simple task such as picture naming (Beatty-Martinez et al., 2020; Casado, Walther, et al., 2023; Yilmaz and Schmid, 2012). On the other hand, when lexical L1 access is compared in the same L2-immersed individuals before and after reimmersion to an L1 environment, there seems to be a difference in lexical L1 access; in particular, after short reimmersion in an L1 environment the L1 lexical access is easier (Baus et al., 2013; Botezatu et al., 2021; Linck et al., 2009). Our behavioural data (Casado, Walther, et al., 2023) demonstrated that when the speed of picture naming in L1 was assessed, Polish-English bilinguals immersed in the L2 environment (United Kingdom) for about 10 years (2 - 24 years) performed similarly as Polish-English bilinguals living in the L1 environment (Poland). Altogether, based on the published studies, as far as picture naming speed is concerned, bilinguals who reside in the L2 environment for a relatively long period, appear to have L1 words equally available as bilinguals who reside in the L1 environment.

Even though no clear detrimental effect of long-term L2 immersion for L1 lexical access has been observed in comparison with speakers in the L1 environment, some studies reported that when L2 immersed speakers are reimmersed in their L1 environment, i.e., revisit their home country for a short time, their speech performance in L1 is facilitated, compared to the time before the visit. This has been shown by faster naming latencies during a blocked picture-naming task in L1 (Casado, Walther, et al., 2023). In particular, in the latter study, we found that Polish-English bilinguals living in the UK showed faster naming latencies in high-frequency words after short reimmersion in Poland, the L1 environment, than during immersion in the UK, the L2 environment. We argued that during their reimmersion in the L1 environment, the speakers primarily encountered high-frequency words, which led to the facilitation of lexical access to these words. This effect was observed even a few days after their return to the L2 environment.

Altogether, based on the behavioural evidence, we can distinguish two distinct aspects of L1 access in migrant bilinguals. On the one hand, it seems that long-term migrants do not necessarily experience difficulties in L1 lexical access compared to bilinguals in their L1 environment (Beatty-Martinez et al., 2020; Casado, Walther, et al., 2023; Yilmaz & Schmid, 2012). On the other hand, the frequency of encountering L1 words or using L1 structures in each language environment impacts the efficiency of accessing L1. Thus, even though long-term immersion in the L2 environment does not necessarily hamper L1 access (at least when assessed via picture naming latencies), migrants show clear sensitivity to changes in language environment: short reimmersion to the native environment appears to temporarily facilitate retrieval of words in L1, especially those that are more frequently used. Still of note is that the assessment of L2 and L1 immersion above was conducted as a between-group comparison of migrants and residents in L1. At the same time, the effects of L1 reimmersion were assessed as a within-group comparison in the migrant population. The conflicting results could therefore result from individual differences and a lack of sensitivity of the used methodology. Considering that previous research exploring L1 access in L2-immersed bilinguals involved uniquely behavioural measures, it is unknown whether we could detect more subtle differences on the neural level, specifically when comparing L2-immersed bilinguals with bilinguals in the L1 environment. As stated above, ERP measures might provide increased sensitivity than behavioural measures even when there is no behavioural change (Bice & Kroll, 2015; McLaughlin et al., 2004) and give us additional insights into the processes underlying bilingual language production.

1.2. Electrophysiological correlates of lexical access

Previous studies identified two ERP components that have shown sensitivity to word retrieval difficulty in the picture-naming task.

The first component is the production P2, characterised by a positive activity peaking around 200 ms in central electrodes (Baus et al., 2020; Strijkers et al., 2010, 2011, 2013). More positive P2 amplitudes have been related to more difficult lexical access. For instance, low-frequency words in L1, which take longer to retrieve than high-frequency words in L1, have been found to evoke more positive amplitudes of the P2 component than high-frequency words (Strijkers et al., 2010). Along the same lines, retrieval of words in L1 with a high level of interlexical competition (evoked by cumulative semantic interference) has been related to more positive amplitudes of the P2 component (Costa et al., 2009). Additionally, L2 words, which in unbalanced bilinguals take longer to retrieve than L1 words, were associated with more positive amplitudes of the P2 component (Strijkers et al., 2013). Altogether, previous studies indicate that the naming P2 component may reflect difficulties in lexical access during word production at the lemma selection stage, according to Indefrey and Levelt's model (2004) of language production.

Another component previously associated with the difficulty of retrieving words during picture naming is the N300. It is characterised by a negative activity peaking around 300 ms in central electrodes. Previously the component has been associated with the recognition of pictures showing less negative amplitudes for non-ambiguous pictures and repeated pictures (Curran et al., 2002; Federmeier & Kutas, 2001; Henson et al., 2004; Philiastides et al., 2006; Schendan & Kutas, 2002, 2007; van Petten et al., 2000). The component is also sensitive to the difficulty of integrating concepts with pictures (Barrett & Rugg, 1990; Eddy et al., 2006; Gratton et al., 2009; Holcomb & Mcpherson, 1994; Philiastides & Sajda, 2006; Voss et al., 2010; West & Holcomb, 2002). Recently, the N300 component was identified as reflecting effort associated with the difficulty of accessing the lexical representation of the to-be-named picture, the task at hand (Wodniecka et al., 2020). More specifically, the more negative N300 amplitudes were observed for naming pictures in more difficult task conditions, e.g. when pictures were named for the first time in the experiment, compared to when they were repeated, or when pictures were named in L1 after a block of previous naming other pictures in L2. That is, the more negative N300 amplitudes seem to reflect difficulties integrating the lexical information associated with the names of corresponding pictures with the perceptual features of the image.

Altogether, we speculate that during picture naming two different components could reflect the difficulty of accessing the picture's name at a different timing; on the one hand, there is the naming P2 around the 200 ms, associated with the lemma selection stage which could be modulated by the amount of competing information active by the lexical concept (Costa et al., 2009; Strijkers et al., 2010; 2013). On the other hand, there is the N300 around 300 ms, associated with the difficulty of integrating the process of picture recognition in synchrony with the word retrieval process of the corresponding name in the target language (Wodniecka et al., 2020). Thus, the N300 component underlies the integration of two subprocesses: visual features that allow the recognition of the picture and the task at hand, which in this case is accessing the picture's name. Therefore, accessing the lexical information (which takes place in the previous time window) is needed to obtain modulations in the following one (N300) corresponding to the integration of the visual and the lexical information (see Valente et al., 2014).

1.3. Current study

In the current study, we explored electrophysiological correlates of lexical access difficulty in native language production associated with (1) long-term immersion in L2; and (2) the effects of short-term reimmersion in the L1. For this aim, we recorded EEG data during a picture naming task in L1 (Polish). The task was part of a large-scale study in which we tested a migrant group of Polish-English bilinguals living in the UK (L2 environment) and a control group of Polish-English bilinguals residing in Poland (L1 environment). Each group was tested twice with a mean between-session interval of 94 days (*SD* = 48, 30 - 259 days). The migrant group was tested once during immersion in the L2 environment and once after short reimmersion in the L1 environment with a variable interval between sessions. The control group was tested both times in the L1 environment, with a similar interval between the tests as the migrant group. The detailed pattern of behavioural data from the picture naming task was already reported in Casado, Walther, et al. (2023). Here we focus solely on the electrophysiological indices of picture naming performance.

We built on previous studies that identified two ERP components during picture naming related to the difficulty of word retrieval during language production: the naming P2 component and the N300 component, both indicators of lexical access difficulty. Based on these results, we expected to observe these components to reflect changes in migrants' language system, also in comparison with control bilinguals, due to changes in the language environment.

We made three predictions. Firstly, given that the ERPs can reflect neural modulations even in the absence of any behavioural effects, we predicted that naming pictures in L1 would be accompanied by an enhanced amplitude of P2 and N300 in the migrant group compared to the control group. Such a result would suggest that the difficulty that migrants experience while retrieving words in L1 is reflected in brain activity during lexical access and its integration with picture recognition. Our second prediction concerned the effects of short-term L1 reimmersion. We predicted reduced

P2 and N300 amplitudes after L1 reimmersion compared to the L2 immersion condition. The third prediction related to the interaction between language environment and frequency of words to be named: we expected that the P2 amplitudes should also be modulated by the lexical frequency of the words; in particular, low-frequency words should evoke more positive amplitudes of the P2 compared with high-frequency words. Finally, a potential differentiation in the effects of P2 and N300 regarding the difference between groups and between before and after L1 reimmersion would give us insight into the time course of the phenomenon during word production. Modulations in the P2 amplitude would indicate changes to the lexical processing steps during word production caused by different language activation levels. On the other hand, modulations in the N300 amplitude could be more closely attributed to the integration of visual and the already active lexical information during picture naming. Therefore, a modulation in the P2 but not in the N300 for the migrant group would indicate that migrants showed greater difficulty accessing L1 words' lexical information, and not a difficulty in integrating visual information during L1 word production, reflected in the later component.

2. **Methods**

2.1. Participants

We analysed data collected from a rather large-scale study in which we tested the consequences of immersion and reimmersion into many language domains by collecting the data from a battery of language-related tasks. In that study, we collected data from various tasks from 55 Polish-English bilinguals living in the United Kingdom (migrant group) and 56 Polish-English bilinguals residing in Poland (control group). Each group was tested twice with a mean between-session interval of 94 days (*SD* = 48, 30 - 259 days). In the present study, we focus on L1 lexical access, assessed through the EEG response to naming pictures in L1, which was available for 32 migrant bilinguals (see details below) and contrasted with a matched sample of 32 control bilinguals.

For the migrant group, we recruited Polish native speakers living in the UK for at least two years. As indicated before, the migrant group was tested twice: "During L2 immersion", that is after at least 30 days fully immersed in the L2 environment, without leaving the UK; and "After L1 reimmersion", after a reimmersion in the L1 environment (Poland), less than 7 days after returning from their L1 environment. From the initial sample of 55, we excluded nine participants who did not complete the session during L2 immersion and another five because they did travel to Poland within 30 days beforehand and did not fulfil the selection criteria. Additionally, nine more participants were excluded due to technical problems while recording the electrophysiological and/or behavioural responses. The final sample included in the analyses consisted of 32 participants.

For the control group, we recruited native speakers of Polish with upper-intermediate English proficiency. Moreover, pre-selection criteria only allowed those participants who spent 30 days before each session in Poland. Like the migrant group, the control group completed two sessions, the "X-Context" session and the "Y-Context" session. In both sessions, they were immersed in the L1 environment, that is, tested in Poland. By having two sessions for the control group as well, we could account for the effects of repetition of the picture naming task within a balanced design. Out of the initial group of 56, seven participants were excluded because they completed only one experimental session, seven additional participants were excluded due to bad quality of the EEG recordings, and three more participants were excluded due to recording errors. The remaining 32 participants were compared to the migrant group.

Table 1

Demographic information and language experience of participants.

	Migrant group (N = 32)		Control group (N = 32)		t-test	
N	32 (29 female)		32 (23 female)			
Participant’s Age (years)	36.16 [33.83, 38.48]		29.84 [27.13, 32.56]		t(62) = -4.19, p = <0.01***	
SES (1 - 10)	6.64 [6.09, 7.20]		5.91 [5.33, 6.48]		t(62) = -1.62, p = 0.11	
Years of education	18.35 [17.43, 19.28]		17.16 [16.39, 17.92]		t(62) = -2.34, p = 0.02*	
Length of residence in L2 environment (years)	9.66 [7.90, 11.41]		-			
Length of reimmersion in L1 environment (days)	13.38 [10.42, 16.33]		-			
Time delay between L1 reimmersion and recording (days)	3.06 [2.39, 3.74]		-			
Self-assessed language experience (1-10)	L1	L2	L1	L2	L1	L2
Self-rated proficiency	9.82 [9.65, 9.99]	7.86 [7.52, 8.19]	9.71 [9.46, 9.96]	7.17 [6.74, 7.60]	t(62) = -0.57, p = 0.57	t(62) = -2.33, p = 0.02*
Speaking	9.64 [9.36, 9.92]	7.68 [7.23, 8.13]	9.56 [9.17, 9.95]	6.59 [6.10, 7.09]	t(62) = -0.20, p = 0.84	t(62) = -3.02, p = <0.01*
Writing	9.77 [9.50, 10.04]	7.42 [6.96, 7.89]	9.66 [9.40, 9.91]	6.66 [6.04, 7.27]	t(62) = -0.28,	t(62) = -1.74,

					p = 0.78	p = 0.09
Listening	9.94 [9.85, 10.02]	7.91 [7.60, 8.22]	9.84 [9.66, 10.03]	7.53 [7.12, 7.94]	t(62) = -0.91,	t(62) = -1.37,
					p = 0.37	p = 0.17
Reading	9.93 [9.85, 10.02]	8.42 [8.07, 8.77]	9.78 [9.50, 10.07]	7.91 [7.43, 8.39]	t(62) = -1.05,	t(62) = -1.74,
					p = 0.30	p = 0.09
Percentage of daily use (%)	40.46 [34.72, 46.21]	59.25 [53.66, 64.85]	81.32 [75.44, 87.19]	16.84 [12.75, 20.93]	t(62) = 10.29,	t(62) = -12.68,
					p = <0.01***	p = <0.01***
Age of L2 acquisition (years)	-	13.05 [11.70, 14.39]	-	10.00 [8.41, 11.59]	t(62) = -3.89, p = <0.01***	
Intensity of language switching (1 = no switching, 10 = always switching)	4.82 [3.90, 5.74]		4.06 [3.27, 4.85]		t(62) = -0.84, p = 0.40	
Objective L2 proficiency measures						
LexTALE (mean accuracy in %)	-	0.78 [0.73, 0.83]	-	0.74 [0.70, 0.77]	t(62) = -1.55, p = 0.13	
General English Test (mean accuracy in %)	-	89.88 [86.44, 93,28]	-	85.12 [82.44, 87.84]	t(62) = -2.54, p = 0.01*	

Note. The first part of the table describes the demographic information of the final migrant group and the final control group. The rows display (1) the number of participants with the number of women in brackets (gender was self-reported by participants), (2) age (in years), (3) socio-economic status on a 1 to 8 scale based on Adler et al. (2000), (4) years of education (in years), (5) length of residence in an L2 environment (in years), (6) length of immersion the L1 environment and (7) time passed between arrival at the airport in the UK and the experimental recording session. The second part of the table summarises the self-assessed language experience based on a questionnaire. The self-rated proficiency is presented on a scale from 1 to

10, where 1 = "no knowledge of a given language" and 10 = "native-like proficiency". The daily use of each language is presented in percentages while the age of acquisition is presented in years. Bilingual switching is presented on a scale from 1 to 10, where 1= "I never switch languages within sentences" and 10 = "I always switch languages within sentences". The objective L2 proficiency measures in English are presented in percentages. 95% confidence intervals are given in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All participants received monetary compensation for participating in the study (65 pounds per session or the equivalent in Polish zloty) and a selection of Polish books. The study met the requirements and received the approval of the Ethics Committee of Jagiellonian University Institute of Psychology concerning experimental studies with human subjects.

All participants learned English as a second language and used it daily (see Table 1), however, the experimental group used L2 significantly more than the control group. We assessed their English proficiency with the General English Test (by Cambridge Assessment: https://www.cambridgeenglish.org/test-your-english/general-english/) and an online version of the LexTALE task (Lemhöfer & Broersma, 2012) programmed in Inquisit (Inquisit 5 [Computer software] (2016). Retrieved from https://www.millisecond.com). The selection criteria for participating in the study were: a self-reported English proficiency of upper-mediate (B2) or above, 70% or higher accuracy in the General English Test, and 60% or higher accuracy in the LexTALE test. The migrant participants also reported that they only used Polish (never English or other languages) when contacting friends and family in Poland. Due to the small target population, we could not recruit a balanced group of migrants concerning gender, age, or language proficiency. Most likely due to the long-term immersion in the L2 environment, the migrant group showed slightly higher L2 proficiency than the control group, even though both of them only demonstrated on average an upper intermediate level of L2. The language experience of both groups is summarised in Table 1.

2.2. Task and procedure

The order of the sessions was counterbalanced for both groups of participants. In the final sample, the session during L2 immersion was the first session for 16 participants of the migrant group and the X-Context for 16 participants of the control group.

2.2.1. Materials

Table 2

Stimuli information for the four subsets used during picture naming.

	A	B	C	D
PL Age of Acquisition (years)	3.43 [3.27, 3.59]	3.56 [3.40, 3.73]	3.41 [3.28, 3.54]	3.49 [3.34, 3.64]
Name agreement (%)	0.93 [0.89, 0.96]	0.92 [0.88, 0.95]	0.95 [0.92, 0,97]	0.93 [0.90, 0.96]
Lexical frequency	2.77 [2.59, 2.95]	2.71 [2.55, 2.88]	2.72 [2.54, 2.89]	2.85 [2.66, 3.04]
Phoneme length	5.58 [5.20, 5.96]	5.38 [4.94, 5.82]	5.72 [5.32, 6.13]	5.60 [5.15, 6.05]

Note. The table shows the average stimuli properties of the used picture subsets. The rows display (1) the age of acquisition for the Polish picture names (in years) based on Haman et al. (2015), (2) the name agreement of the used pictures (in percent) based on Wolna et al. (2022), (3) lexical frequency based on Mandera et al. (2015), (4) phoneme length and (5) Imageability based on Wolna et al. (2022). 95% confidence intervals are given in parentheses.

The stimuli included in the picture-naming task consisted of 216 coloured images from the Cross-Linguistic Lexical Tasks database (Haman et al., 2017). We divided all pictures into four subsets. These subsets of pictures were balanced for name agreement (based on Wolna et al., 2022), lexical frequency (based on Mandera et al., 2015), age of acquisition (Haman et al., 2015), and mean length in phonemes. Polish-English cognates and homophones were excluded during our stimuli selection process. Moreover, each subset contained a comparable number of images from different semantic categories. The four subsets were counterbalanced across participants and sessions so that no pictures were repeated between sessions and each participant saw two out of four picture subsets.

The variation in the lexical frequency of pictures' target names allowed us to explore possible interactions between the L1 vs. L2 environments and the lexical frequency.

2.2.2. Procedure

In the picture-naming task, pictures were displayed on a computer screen using DMDX (Forster & Forster, 2003). The pictures were presented in the centre of the screen. Each trial was preceded by a black screen presented for 1000 ms, followed by a fixation cross in the screen's centre for 1000 ms. A picture was then shown in the centre of the screen until the participant responded or until the time to respond was over (3000 ms). The participants were instructed to name pictures aloud in their native language as quickly and accurately as possible. Each session of picture naming had a total of 58 trials (4 practice trials and 54 regular trials). Overall, the picture-naming task lasted approximately 5 minutes.

2.2.3 EEG Recording

The EEG was recorded during the picture-naming task at 1024 Hz from 32 Ag/AgCl scalp electrodes positioned at the standard 10-20 locations, mounted in an elastic cap, using the Biosemi Active Two recording system. Electrodes were initially referenced online to the Common Mode Sense electrode located at the C1 electrode. The horizontal and vertical electrooculogram (EOG) was recorded bipolarly using electrodes placed below and above a participant's left eye and at the outer canthus of each eye, respectively. The EEG signal was offline filtered with a band-pass filter (0.1 – 25 Hz frequency range; low cutoff slope: 24 dB/oct; high cutoff slope: 12 dB/oct) and re-referenced offline to the mean of the left and right mastoids. The data was offline preprocessed using BrainVision Analyzer (Brain Products, Gilching, Germany), downsampled to 256 Hz, and baseline-corrected. We extracted segments of 900 ms (-100 to 800). Ocular artefacts were removed with Independent Component Analysis (ICA, (Delormeet al., 2007; Jung et al., 2000)) by calculating the ICA components based on 1 Hz high-pass filtered data and removing the artifactual components. On average 2.8 components were removed per participant with a maximum of 6 components being removed. The corrected ICA matrix was applied to the 0.1 Hz high-pass filtered data set. Segments containing artefacts were cleaned manually. Afterward, the data was exported to Matlab for further analysis using EEGlab (Delorme & Makes, 2004) and ERPlab (Lopez-Calderon & Luck, 2014). Mean amplitudes were calculated with specific time windows and electrode selections for each ERP component, as described below.

2.3. Analysis

Two time-windows were selected for the EEG analysis based on previous studies (Costa et al., 2009; Strijkers et al., 2010, 2011, 2013; Wodniecka et al., 2020): the naming P2 around 175 ms (160 – 240 ms) and the N300 around 300 ms (250 – 350 ms). For each time window, the mean amplitude was calculated with a different selection of electrodes following the literature: FC1, FC2, Fz, Cz, CP1, CP2 for the naming P2 (Strijkers et al., 2013), and FC1, FC2, C3, C4, Cz, CP1, CP2 for the N300 (Wodniecka et al., 2020).

To be able to relate the electrophysiological results of the present study with the findings of the behavioural analysis already reported in Casado, Walther, et al. (2023), we included all variables we controlled for in the previous study in the analyses to maintain the two analyses as similar as possible. That is, besides the main interest factors (Group, Context, and Word-lexical frequency), we included some of the participants' variables the groups were not matched for: the participants' age, and the age of L2 acquisition. Moreover, given that there were differences in L2 proficiency between groups, we also included this variable as a covariate in the main model.

Responses with naming latencies below 650 ms were removed from the data to avoid artefacts in the signal evoked by the articulatory movements (9.95%). We also excluded trials with inaccurate responses (4.24%). In total, 14.19 % of the trials were excluded.

Table 3

Summary of the final trial count after artefact rejection during EEG processing and removal of inaccurate responses and too early articulator movements.

	During L2 immersion/X-Context	After L1 reimmersion/Y-Context

Migrant group	88.6% (1531 trials)	82.0% (1417 trials)
Control group	87.0% (1504 trials)	85.6% (1479 trials)

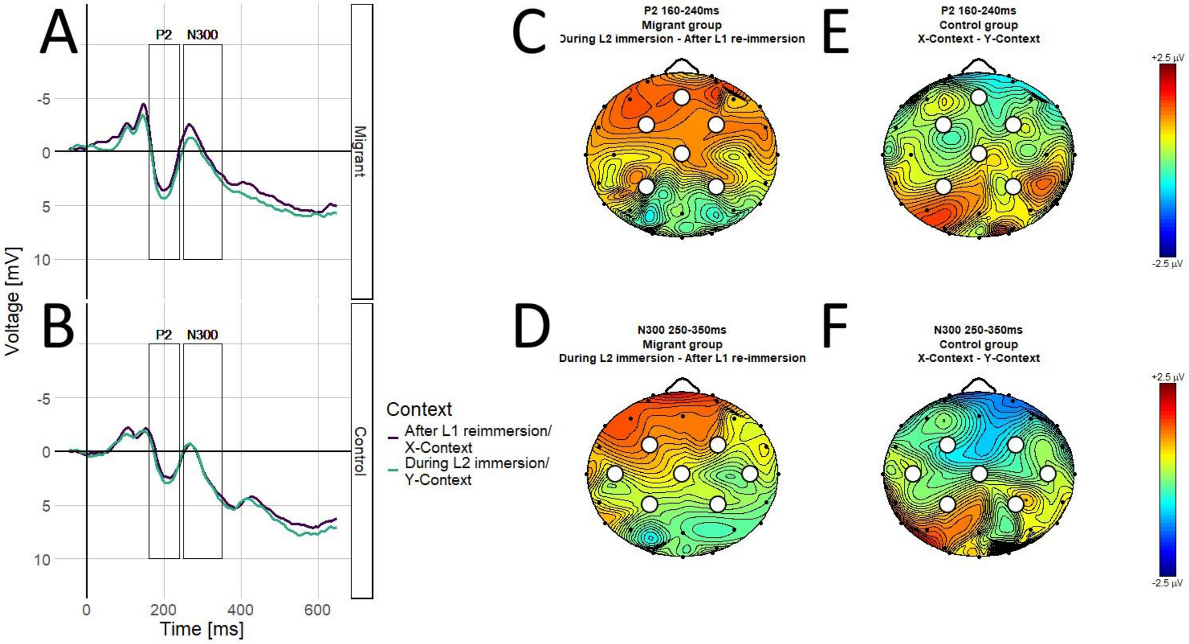
Note. The table gives the percentage and number (in brackets) of remaining trials in the final analysis for each Group and Context. The raw data set contained 1728 trials within each condition.

We used R Studio (R Development Core Team, 2020, Version 4.0.2) to fit linear mixed-effects models with the lme4 library (Bates et al., 2015, Version 1.1-23). The general models included the mean amplitude of P2 and N300 as the dependent variables and participants and pictures as crossed random effects. As fixed effects, we included Group (Control, Migrant), Context (During L2 immersion/X-Context, After L1 reimmersion/Y-Context), Word-lexical frequency (based on Mandera et al., 2015), participant's Age, Age of L2 acquisition and L2 proficiency (based on the mean scores of the LexTALE and Cambridge proficiency tasks), and the Trial number. We also included the interactions between Group, Context, and Word-lexical frequency. Before running the analyses, all categorical predictors were deviation-coded using a sum contrast (Group: Control group = -0.5, Migrant group = 0.5; Context: During L2 immersion/X-Context = -0.5, After L1 reimmersion/Y-Context = 0.5; Session order: first session = -0.5, second session = 0.5). The continuous predictors of Word-lexical frequency, participant's Age, Age of L2 acquisition, and L2 proficiency were centred and scaled. The Trial number was log-transformed. The maximal model also included a by-picture random intercept and random slopes for Group, Context, participant's Age, Age of L2 acquisition, and L2 proficiency. Additionally, we used a by-Participant random intercept with random slopes for Context and Word-frequency.

We fitted the maximal models first (Barr et al., 2013). If the model did not converge, we first removed correlations between random effects and in the next step, the random effects with the smallest unique variance, following the recommendation by Bates et al. (2018). Summaries and p-values of the models were calculated using the lmerTest package (Kuznetsova et al., 2017). Absolute t-values greater than two were considered significant. For pairwise comparisons, the emmeans and lsmeans packages were used (Lenth, 2016, 2021).

3. Results

Figure 1



Note. Stimulus-locked grand average ERP waveforms for the Migrant **(A)** and Control **(B)** groups and scalp topography maps for naming P2 **(C, E)** and N300 **(D, F)**. Waveforms depict mean voltages averaged over the shared electrodes between the P2 and N300 clusters: FC1, FC2, Fz, Cz, CP1, and CP2. Scalp plots show differences in voltage between different language contexts for the migrant group and between the two testing sessions for the control group. Selected electrode clusters used for calculating mean amplitudes are marked in white.

The results are presented in Table 4 and Table 5. The detailed analysis of the behavioural data was reported in Casado, Walther, et al. (2023). Additionally, a summary is presented in Table 6. In the P2 component analysis, there was no main effect of Group. Still, we found a significant interaction between Group and Context, demonstrating a modulation of the P2 component in the migrant group due to the language context: more positive P2 amplitudes were observed during L2 immersion compared to after L1 reimmersion (see Figure 1 panel A). As expected, no differences in the P2 amplitudes between contexts were found for the control group (see Table 6). However, it should be noted that the distribution of the P2 was more frontal than expected based on the previous studies (Costa et al., 2009; Strijkers et al., 2010; 2011; especially Strijkers et al., 2013).

Contrary to our predictions, we did not observe the effect of frequency on mean amplitude in the P2 component. Additionally, there was a significant effect of the participant's Age on the P2 amplitude with an increased amplitude for older participants.

In the analysis of the N300 component, we found no main effects of Group or Context (see Figure 1 panel B) and no significant interactions.

Table 4

Fixed effects for the LME model with ERP amplitude and both groups for the P2 component.

	Estimate	SE	t	by- Picture SD	by- Participant SD
Intercept	-0.03	0.08	-0.44	0.10	0.31

Group	-0.05	0.10	-0.47		-
Context	0.03	0.03	0.86		0.16
Word-lexical frequency	0.00	0.01	0.22	-	
Participant’s Age	0.22	0.05	4.71***		-
Age of L2 acquisition	-0.09	0.05	-1.91		-
L2 proficiency	-0.07	0.04	-1.58		-
log (Trial number)	0.01	0.02	0.54		
Group: Context	0.18	0.06	2.76**		-
Group: Word-lexical frequency	0.00	0.02	-0.12	-	-
Control Group: Context: Word-lexical frequency	0.00	0.03	0.08	-	-
Migrant Group: Context: Word-lexical frequency	0.03	0.03	0.89	-	-

Table 5

Fixed effects for the LME model with ERP amplitude and both groups for the N300 component.

	Estimate	SE	t	by- Picture SD	by- Participant SD
Intercept	-0.09	0.09	-0.98	0.23	0.39
Group	-0.18	0.12	-1.50		-
Context	0.02	0.04	0.63		0.20
Word-lexical frequency	0.02	0.02	1.15	-	
Participant’s Age	0.06	0.06	1.10		-
Age of L2 acquisition	-0.07	0.06	-1.26		-
L2 proficiency	0.01	0.05	0.21		-
log (Trial number)	0.03	0.02	1.35		
Group: Context	0.11	0.07	1.52		-
Group: Word-lexical frequency	0.00	0.02	-0.05	-	-
Control Group: Context: Word-lexical frequency	0.00	0.03	-0.10	-	-
Migrant Group: Context: Word-lexical frequency	0.02	0.03	0.61	-	-

Table 6

Summary of the electrophysiological data.

ERP component	Group	During L2 immersion / X-Context	After L1 reimmersion / Y-Context	Context comparison
P2	Migrant group	0.04 (0.07)	-0.08 (0.06)	*z* = -2.59; *p* = 0.01*
	Control group	-0.01 (0.07)	0.05 (0.06)	*z* = 1.33; *p* = 0.19
	Group comparison	*z* = -0.40; *p* = 0.69	*z* = 1.40; *p* = 0.16	
N300	Migrant group	-0.04 (0.08)	-0.11 (0.08)	*z* = -1.55; *p* = 0.12
	Control group	0.08 (0.08)	0.11 (0.08)	*z* = 0.62; *p* = 0.53
	Group comparison	*z* = 1.00; *p* = 0.32	*z* = 1.87; *p* = 0.06	

Note. The table gives the predicted amplitudes of the two ERP time windows for the two groups in the two Contexts and the pairwise comparison from their respective linear-mixed models using the emmeans function. Standard errors are given in parentheses.

4. Discussion

In the present study, we explored neural signatures of L1 production by bilinguals in an L1 environment (native language country) and an L2 environment (foreign language country). To this aim, we compared the EEG response evoked during an L1 picture-naming task of Polish-English migrants immersed in the L2 environment with that of Polish-English bilinguals in the L1 environment. Moreover, we explored the migrants' L1 lexical access under two different conditions: during L2 immersion, and after short-term reimmersion in the L1 environment. We focused on two ERP components: P2 and N300, previously associated with the ease of lexical access during language production (Strijkers et al., 2010; 2011; Wodniecka et al., 2020).

We tested three predictions. First, if long-term immersion in the L2 environment results in decreased accessibility of L1 words, we should observe larger amplitudes of the P2 (more positive) and N300 (more negative) effects in the migrant group compared to the control group. Second, if short-term reimmersion in L1 can reverse the decreased access to L1 words due to long-term immersion in L2, we should observe lower amplitudes of the P2 (less positive) and N300 (less negative) components in migrants after short reimmersion in the L1 context compared to migrants during L2 immersion. Finally, if the decreased access to L1 words primarily affects high-frequency words (as we argue in Casado, Walther, et al., 2023), we should observe larger differences in frequencies for P2 amplitudes in the migrant group after reimmersion in L1 than during immersion in L2. Moreover, we wanted to use the potential differentiation in the effects of P2 and N300 components to shed some light on the time course with which immersion and reimmersion modulate the word production process. If difficulties in picture naming arise during the lexical access stage, we should observe modulations in the naming P2 component. In contrast, if difficulties in picture naming arise during the integration of the lexical information with the perceptual features of the picture, we should observe a modulation in the N300 component.

Regarding the first prediction, the results of the present analyses showed no ERP modulation related to overall differences between migrants and control bilinguals. The absence of the overall group effect in the ERPs mirrors the lack of overall differences between the groups in L1 naming latencies (see Footnote 5 and Casado, Walther, et al., 2023). We, therefore, conclude that despite the long-term immersion of the Polish-English migrants living in the L2 environment (in the UK), we did not find evidence for reduced access to the native language, either in behavioural responses or in the electrophysiological markers of the lexical access. Still, the P2 component was sensitive to short-term changes in migrants' language environments, confirming our second prediction. That is, more positive P2 amplitudes were evoked during L2 immersion vs. after L1 reimmersion, indicating that short-term reimmersion in the L1 environment resulted in neurocognitive adaptations as reflected in the modulation of the early frontal positivity. Thus, we conclude that the neurocognitive adaptations occur quite early in the process of retrieving the picture's name. It is unclear, however, whether the observed positivity can be identified as the previously reported P2 component (Costa et al., 2009; Strijkers et al., 2010; 2013), as its topographic distribution seems to be more frontal. Finally, we found no evidence for the third prediction, as there was no effect of the lexical frequency; that is, the P2 was insensitive towards any frequency manipulation. The N300 component was also not modulated by any of the variables (migrants vs. control; migrants' L2 immersion vs. L1 reimmersion), thus showing that the effects of immersion and reimmersion during word production do not affect the integration process of the visual features of the picture and the lexical information corresponding to the name.

Below, we discuss the implications of our findings for our understanding of the long- and short-term adaptations of bilinguals' cognitive system to the language environment.

4.1. Effects of long-term immersion

In line with our behavioural results reported in Casado, Walther et al. (2023), we observed no significant differences in either P2 or N300 mean amplitudes when comparing the evoked responses of migrants with those of the control group. This pattern of results also aligns with previous behavioural results which found that the lexical access in the native language of migrants living in L2 environments for a long time does not differ from bilinguals living in an L1 environment (Beatty-Martinez et al., 2020; Yilmaz & Schmid, 2012) or an environment in which two languages are frequently mixed (Beatty-Martinez et al., 2020). Together with the data reported here, the evidence suggests that long-term immersion in the L2 does not necessarily result in generalised difficulty in accessing native language words. Even after long-term L2 immersion, migrants can maintain similar L1 levels as bilinguals who live in an L1 environment.

The fact that access to L1 lexical information is not necessarily negatively affected by immersion in the L2 environment (as shown by the absence of group differences in naming latencies and in mean amplitudes) is consistent with the interface hypothesis originally proposed by Sorace and Filiaci (2006). Although their framework explains the difficulties experienced by highly proficient balanced bilinguals, it can also relate to the unbalanced bilinguals with intermediate proficiency, like the participants in the current study. Sorace and Filiaci posited that long-term L2 immersion only affects the sensitivity to high-level structures external to grammar, such as the interface between syntax and pragmatics that determine appropriateness in context (e.g., during anaphora resolution, see Chamorro et al., 2016; Chamorro & Sorace, 2016; Köpke & Genevska-Hanke, 2018), rather than causing a permanent change in speakers L1 knowledge representations like vocabulary (Chamorro & Sorace, 2019; Sorace, 2011; 2016). Here we provide the first piece of evidence of electrophysiological investigation suggesting that low-level linguistic structures related to the formal semantic features internal to grammatical representations (i.e., lexical level) indeed do not seem to be affected by L2 immersion. Thus, considering that the lexical information is part of the lexicon and that we did not find differences in the lexical access abilities between immersed and non-immersed bilinguals, we would propose that long-term L2 immersion does not necessarily modulate the L1 lexicon permanently. In short, these results add to the accumulated body of research indicating that difficulties accessing L1 encountered by bilinguals long-term immersed in the L2 environment affect online sensitivity rather than causing a permanent change in the speaker's L1 knowledge representations (Chamorro et al. 2016; Sorace, 2011; 2016).

4.2. Short-term changes triggered by the language environment: L2 immersion vs. L1 reimmersion

Concerning the short-term manipulation of the migrants' language environments, we found that the short-term reimmersion in the L1 environment was related to a change in neural response to picture naming in L1. More specifically, after short-term L1 reimmersion there was a reduction of the positivity at the 160 - 240 ms time window in frontal regions. We intended to explore modulations of the naming P2 component, however previous literature characterised the component slightly differently from the one we found. In particular, the P2 component was previously associated with the ease of lexical access with more positive amplitudes under more difficult conditions (Costa et al., 2009; Strijkers et al., 2010, 2011). Following this interpretation, the decreased amplitude of P2 after L1 reimmersion could indicate that migrants benefited from spending time in the L1 environment, such as this short visit to the native language country facilitated access to L1 compared to during immersion in the L2 environment. However, the interpretation of the P2 component as an index of ease of lexical access was based on its sensitivity to word-lexical frequency, such as low-frequency words evoked more positive P2 amplitudes than high-frequency words (Strijkers et al., 2010). Yet, in the present study, we did not find a modulation of the positivity at the 160 - 240 ms time window due to the word-lexical frequency, either as a main effect (which could be expected based on Strijkers et al., 2010) or in an interaction with the environment (unlike in the behavioural analyses reported in Casado, Walther, et al., 2023). Therefore, it is unclear whether the positivity observed in our study is the naming P2 identified in previous research and if it could be interpreted as an index of lexical access. As indicated above, the overall distribution of our frontocentral component differs from the typical broad centroparietal distribution reported in previous studies characterising the P2 as a lexical access index (e.g., Costa et al., 2009; Strijkers et al., 2010; 2011) (see Figure 1 panel C). Instead, the distribution coincides with the more frontal component reported in Strijkers al. (2013). Strijkers et al. found a different distribution from the one observed in the current study for the language effect (central, in L1 vs. L2 comparison) and the lexical frequency (centroparietal, for low vs. high-frequency word comparison). Therefore, it is likely that the positivity in the 160 - 220 ms time window found in the present study does not reflect the difficulty of lexical access, but probably a different process, to be better defined. One previous study using a picture-naming task reported a P2 component similar in distribution —frontocentral— to the one reported here (Branzi et al., 2014) but offering a different interpretation of the P2 than Strijkers et al.'s (2010). Rather than seeing the P2 component as an index of lexical access difficulty, Branzi et al. (2014) suggested that this P2 component could reflect the mechanisms of language control; in particular, it could index the application of proactive control. In their study, Branzi et al. (2014) investigated the scope of global-local language control in a group of balanced bilinguals. The participants named new and repeated pictures in their native or second language in blocks divided by language following different orders: L1-L2-L1 or L2-L1-L2. Under the L1-L2-L1 order, the authors found that when balanced bilinguals named pictures in L1 after using L2, more positive P2 amplitudes were evoked compared to the first block of L1 naming, for both, repeated and new items. The authors argued that the P2 component in their study reflected language control mechanisms applied during L1 lexical access to manage the persisting activation of the previously used language proactively —in their design, L2— that would create interference within the subsequent naming in the L1.

Following the interpretation of the P2 component proposed by Branzi et al. (2014), our results may indicate that short-term reimmersion in L1 influences the proactive control mechanisms applied to control for language interference during L1 lexical access. In our study, we found a modulation of the positivity in the 160 - 220 ms time-window of the migrant group under the different language environments: the positive amplitudes were decreased after a reimmersion in the L1 environment compared to during immersion in the L2 environment. In our case, the higher positive amplitudes during immersion in the L2 environment could be reflecting the application of proactive control to regulate the interference from the L2 when accessing L1. Similarly, the less positive P2 amplitudes after reimmersion in the L1 environment could reflect a decrease in the application of proactive control during L1 access, given the reduced interference from L2 in the L1 environment.

Two other studies support the interpretation suggesting that proactive control plays a crucial role in regulating interference between L1 and L2 in different language environments. Beatty-Martínez et al. (2020) and Zhang et al. (2021) both demonstrated that bilinguals living in their L2 environment, compared to bilinguals in the L1 environment, are more likely to use proactive strategies for cognitive control, measured with the AX-CPT task (Braver et al., 2007). Both studies argue that constant exposure to L2 in the L2 environment trains bilinguals to use alternative control strategies to fight interference from the unwanted language (L2 when accessing L1), like proactive control. This interpretation can also be reconciled with the discrepancies between two different P2 distributions observed in comparisons between languages (Strijker et al., 2013) and in comparisons between low- and high-frequency words (Baus et al., 2020; Strijkers et al., 2010, 2011). While both these effects have been previously proposed to reflect lexical access difficulty, they may correspond to two different mechanisms, both relevant to bilingual speech production. The P2 component found in centroparietal electrodes (Baus et al., 2020; Strijkers et al., 2010) can reflect lexical access difficulty that drives the word-frequency effect. On the contrary, the P2 component found over the frontocentral electrodes may reflect the engagement of proactive control. Engagement of proactive control due to increased interference between languages can be well justified in all studies that report the frontocentral P2. First, in Strijkers et al. (2013), a more positive P2 amplitude is observed in speech production in L2, a weaker language that needs to deal with interference from the more strongly activated L1. Second, Branzi et al., (2014) report a more positive frontocentral P2 in response to naming in L1 after L2 compared to a baseline L1 naming. In this case, the use of L2 before L1 can lead to a temporary increase in the L2 activation level that results in increased interference between the two languages in the subsequent production of L1 (for the discussion of the influence of recent language use on bilingual language control see Casado, Walther, et al., 2023; Bialystok, 2024; Blanco-Elorrieta & Caramazza, 2021). Finally, our results show more positive frontocentral positivity (P2) in response to speech production during long-term immersion in L2 compared to a context of short-term reimmersion in L1. As previously explained, living in the L2 environment may increase L2 activation and in consequence, it may lead to higher interference between languages.

4.3. The time course of the role of immersion and reimmersion during word production

We decided to explore the N300 component as it was previously identified as a marker of L1 lexical retrieval difficulties during picture recognition (Wodniecka et al., 2020) and could give us information about whether the possible difficulties in L1 access could be attributed to the integration of visual and lexical information during picture naming at the different environments. We did not find modulations of the mean amplitude of the N300 component, neither in the between-group comparison nor when comparing migrants during L2 immersion and after L1 reimmersion. The lack of N300 modulation in the between-group comparison aligns with the behavioural results (Casado, Walther, et al., 2023), confirming that L1 retrieval abilities during picture recognition were not affected in migrants because of long-term immersion in the L2 environment. Moreover, the lack of modulation by the context in the migrant group (L2 immersion vs. L1 reimmersion) supports the idea that the positivity between 160 and 240 ms (P2 component) found in this study was not capturing the ease of lexical access per se, but could have been related to the application of proactive language control during L2 immersion vs. after L1 reimmersion. In sum, the lack of N300 modulations indicates that the changes in the environment did not induce difficulties in the integration during picture recognition.

Altogether, these results could indicate that bilinguals adapt to different language environments by adjusting the language control mechanisms, probably during the lexical processing steps of word production (Indefrey & Levelt, 2004). We performed an exploratory analysis in which we tested the interaction between Group and Context over a wide range of time windows and distributions to test this hypothesis (see Appendix 6.1). We observed a more widespread scalp distribution and a longer duration of the ERP modulation due to L1 reimmersion. This may indicate that changes in the language environment for the migrant group can be traced through many processes involved during word production and not only related to lexical access. Still, additional studies are required to explore this finding and its functional significance more deeply.

4.4. Other relevant findings

In the P2 analysis, we observed a main effect of age, such that the older the participants, the harder it was to name the pictures as reflected by more positive P2 amplitudes. This result may be explained by two possible sources. On the one hand, older participants experience a decline in domain-general mechanisms including slower processing speed, lower working memory capacity, and decreased inhibitory control (Cabeza et al., 2018; Salthouse, 2010). These aspects also affect specific language processes including word retrieval, which has been shown to decrease RT and increase errors (Burke & Shafto, 2008; Peelle, 2019). On the other hand, previous research showed that older participants have greater vocabulary knowledge (Park et al., 2002; Verhaeghen, 2003), which is related to increased interference due to enriched semantic information in older people (Buchler & Reder, 2007; Ramscar et al., 2017). Given that we did not include cognitive tests in our testing battery, we cannot rule out any of the possibilities and further research is needed to clarify it.

Apart from chronological age, we also included L2 proficiency in the analyses as a control variable given that we could not match the participants' group by the General English Test despite our efforts. Our analyses showed that the L2 proficiency did not modulate the ERP response in L1 picture naming. Despite in word recognition studies the modulation of L2 proficiency is expected (e.g., Marian et al, 2003; Spivey & Marian, 1999), it does not seem to be such a clear case for production studies. For instance, Klaus et al. (2018) designed a picture-word interference paradigm in L1 with distractors L2 words presented auditorily, which could be either the translation equivalent of the L1 picture name or unrelated words. They found increased naming latencies when the distractor was the translation equivalent of the to-be-named picture compared to an unrelated distractor, demonstrating that even when speaking in one's dominant L1, translations from the less dominant L2 were active. Importantly, the interference effect was not modulated by L2 proficiency levels. Thus, the naming latencies in L1 did not depend on L2 proficiency. They argue that the L2 proficiency of their participants was rather similar, as in our participants' sample.

Further studies should be conducted including a wider L2 proficiency range to explore whether differences in proficiency could modulate L1 naming latencies

5. Conclusion

In this study, we showed that migrants adapt their neurocognitive system depending on the language environment they are immersed. We observed modulations in the electrophysiological response to naming pictures in the native language, such as more positive amplitudes during L2 immersion and more negative amplitudes after L1 reimmersion in frontal electrodes during the 160 - 240 ms time window (P2 component). Following Branzi et al. (2014)'s interpretation of the P2 component as a marker of proactive control applied by bilinguals to regulate language interference and aligning with Beatty-Martínez et al. (2020) and Zhang et al. (2021) behavioural studies, we propose that the modulation of the early positivity observed in our study reflects the engagement of proactive control in speech production. Under this explanation, the amount of proactive control applied during immersion in the L2 environment would be higher as the level of L2 activation in this context may be increased. Consequently, controlling the interference between languages in the L2 environment may be more effortful than in the L1 environment. Moreover, the amount of proactive control applied to control for between-language interference would be reduced following short-term L1 reimmersion. Altogether, our results could be taken as evidence for the hypothesis that the bilingual cognitive system adapts to language environments, adjusting strategies of language control as more or less proactive, depending on the current needs and providing compensatory mechanisms to mitigate the hindering effects of L2 environment on L1 lexical access.

6. Limitations and future directions

The current study is limited by its methodological constraints. Despite trying our best to match the control group with the chronological age, years of education, and L2 proficiency of the migrant group, some differences proved unavoidable, and we could only account for them statistically in the analyses. Moreover, the two groups were tested in different laboratories (one at the University of Edinburgh, and one at the Jagiellonian University in Krakow), which may have induced environmental and technical differences despite the close similarity between labs and equipment. Furthermore, despite our attempts to control for the time passed between their return from holidays in Poland and their testing session "L1 reimmersion", there was some variability between participants, such that some participants were tested the same day and others up to 7 days after. In the latter case, the participants had already been re-exposed to the L2 environment, which might have attenuated some of the effects of L2 reimmersion.

Future studies should replicate the findings with other populations; in particular, it is to be noted that the migrants who participated in our study maintained close contact with the Polish community and culture, as indicated by the relatively high intensity of L1 use. The contact was also maintained by regular trips to their homeland which allowed us to implement the current design. Possibly, this close contact may have induced a protection against reducing the ability to access the native language, which may vary in different populations.

Authors' note

The data used in the presented models is available here: https://osf.io/bfkp5/?view_only=faab088a3ecd4d9981d3e9e6e636893d.

The authors declare that they have no conflict of interest.

Acknowledgments

The research was possible thanks to a National Science Centre grant (2015/18/E/HS6/00428) awarded to Z.W. and PID2022-142345N-I00 from the Spanish Ministry of Science and Innovation from which A.C. is funded. Special thanks to Dorota Peszkowska and Katarzyna Przybycień, who helped with participant recruitment and testing in Edinburgh. The authors gratefully acknowledge the help of all members of the Psychology of Language and Bilingualism Laboratory, LangUsta, who contributed to the research project by collecting and coding the data, and Michael Timberlake for proofreading.

7. References

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. Health Psychology, 19(6), 586.

Ammerlaan, T. (1996). 'You get a bit wobbly...': Exploring bilingual lexical retrieval processes in the context of first language attrition. [Sl: sn].

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models; 2015. Preprint.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. Acta Psychologica, 142(3), 402–409. https://doi.org/10.1016/j.actpsy.2013.01.010

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. Journal of Experimental Psychology. Learning, Memory, and Cognition, 46(6), 1022–1047. https://doi.org/10.1037/xlm0000770

Beatty-Martínez, A. L., & Titone, D. A. (2021). The quest for signals in noise: Leveraging experiential variation to identify bilingual phenotypes. Languages, 6(4), 168.

Bialystok, E. (2024). Bilingualism modifies cognition through adaptation, not transfer. Trends in Cognitive Sciences.

Bice, K., & Kroll, J. F. (2015). Native language change during early stages of second language learning. NeuroReport, 26(16), 966.

Blanco-Elorrieta, E., & Caramazza, A. (2021). A common selection mechanism at each linguistic level in bilingual and monolingual language production. Cognition, 104625. https://doi.org/10.1016/j.cognition.2021.104625

Bogulski, C. A., Bice, K., & Kroll, J. F. (2019). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. Bilingualism: Language and Cognition, 22(5), 1052-1067.

Botezatu, M. R., Kroll, J. F., Trachsel, M. I., & Guo, T. (2021). Second language immersion impacts native language lexical production and comprehension. Linguistic Approaches to Bilingualism.

Branzi, F. M., Martin, C. D., Abutalebi, J., & Costa, A. (2014). The after-effects of bilingual language production. Neuropsychologia, 52, 102-116.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. Variation in working memory, 75, 106.

Buchler, N. E., & Reder, L. M. (2007). Modeling age-related memory deficits: a two-parameter solution. Psychology and aging, 22(1), 104.

Burke, D. M., & Shafto, M. A. (2004). Aging and language production. Current directions in psychological science, 13(1), 21-24.

Cabeza, R., Albert, M., Belleville, S., Craik, F. I., Duarte, A., Grady, C. L., ... & Rajah, M. N. (2018). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. Nature Reviews Neuroscience, 19(11), 701-710.

Casado, A., Walther, J., Wolna, A.,-Szewczyk, J. , Sorace, A., & Wodniecka, Z. (2022). Advantages of coming back: how reimmersion in the native country impacts migrants' native language access. [manuscript in review] Institute of Psychology, Jagiellonian University.

Chamorro, G., Sorace, A., & Sturt, P. (2016). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition*. Bilingualism: Language and Cognition, 19(3), 520–532. https://doi.org/10.1017/S1366728915000152

Chamorro, G., & Sorace, A. (2019). The Interface Hypothesis as a framework for studying L1 attrition. In The Oxford handbook of language attrition (pp. 24-35). Oxford University Press.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical Selection in Bilinguals: Do Words in the Bilingual's Two Lexicons Compete for Selection? Journal of Memory and Language, 41(3), 365–397. https://doi.org/10.1006/jmla.1999.2651

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. Proceedings of the National Academy of Sciences, 106(50), 21442-21446.

Delorme A & Makeig S (2004) EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics, Journal of Neuroscience Methods 134:9-21

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage, 34(4), 1443-1449.

Ecke, P., & Hall, C. J. (2013). Tracking tip-of-the-tongue states in a multilingual speaker: Evidence of attrition or instability in lexical systems? International Journal of Bilingualism, 17(6), 734–751. https://doi.org/10.1177/1367006912454623

Finkbeiner, M., Almeida, J., Janssen, N., & Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(5), 1075–1089. https://doi.org/10.1037/0278-7393.32.5.1075

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. Behavior Research Methods, Instruments, & Computers, 35(1), 116–124. https://doi.org/10.3758/BF03195503

Gargiulo, C., & van de Weijer, J. (2020). Anaphora resolution in L1 Italian in a Swedish-speaking environment before and after L1 re-immersion: A study on attrition. Lingua, 233, 102746. https://doi.org/10.1016/j.lingua.2019.102746

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. Journal of Memory and Language, 58(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. Bilingualism: Language and Cognition, 1(2), 67–81. https://doi.org/10.1017/S1366728998000133

Haman, E., Łuniewska, M., Hansen, P., Simonsen, H. G., Chiat, S., Bjekić, J., … Armon-Lotem, S. (2017). Noun and verb knowledge in monolingual preschool children across 17 languages: Data from Cross-linguistic Lexical Tasks (LITMUS-CLT). Clinical Linguistics & Phonetics, 31(11–12), 818–843. https://doi.org/10.1080/02699206.2017.1308553

Haman, E., Łuniewska, M., & Pomiechowska, B. (2015). 8. Designing Cross-Linguistic Lexical Tasks (CLTs) for Bilingual Preschool Children. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), Assessing Multilingual Children (pp. 196–240). Multilingual Matters. https://doi.org/doi:10.21832/9781783093137-010

Hulsen, M. E. H. (2000). Language loss and language processing: Three generations of Dutch migrants in New Zealand. [S.l. : s.n.]. Retrieved from https://repository.ubn.ru.nl/handle/2066/18901

Indefrey, P., and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144. doi: 10.1016/j.cognition.2002.06.001

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clinical Neurophysiology, 111(10), 1745-1758.

Klaus, J., Lemhöfer, K., & Schriefers, H. (2018). The second language interferes with picture naming in the first language: Evidence for L2 activation during L1 production. Language, Cognition and Neuroscience, 33(7), 867-877.

Köpke, B., & Genevska-Hanke, D. (2018). First Language Attrition and Dominance: Same or Different? Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.01963

Köpke, B., & Schmid, M. S. (2004). Language attrition. First Language Attrition: Interdisciplinary Perspectives on Methodological Issues, 28(1).

Kuznetsova A, Brockhoff PB, Christensen RHB (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, **82**(13), 1–26.doi:10.18637/jss.v082.i13.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. Behavior research methods, 44(2), 325-343.

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. Journal of Statistical Software, 69(1), 1–33. https://doi.org/10.18637/jss.v069.i01

Lenth, R. V. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. https://CRAN.R-project.org/package=emmeans

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing Access to the Native Language While Immersed in a Second Language: Evidence for the Role of Inhibition in Second-Language Learning. Psychological Science, 20(12), 1507–1515. https://doi.org/10.1111/j.1467-9280.2009.02480.x

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. Frontiers in human neuroscience, 8, 213.

Mandera, P., Keuleers, E., Wodniecka, Z., & Brysbaert, M. (2015). Subtlex-pl: Subtitle-based word frequency estimates for Polish. Behavior Research Methods, 47(2), 471–483. https://doi.org/10.3758/s13428-014-0489-4

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. Brain and language, 86(1), 70-82.

Park, D. C. (2002). Aging, cognition, and culture: a neuroscientific perspective. Neuroscience & Biobehavioral Reviews, 26(7), 859-867.

Peelle, J. E. (2019). Language and aging. The Oxford handbook of neurolinguistics, 10, 1-21.

Ramscar, M., Sun, C. C., Hendrix, P., & Baayen, H. (2017). The mismeasurement of mind: Life-span changes in paired-associate-learning scores reflect the "cost" of learning, not cognitive decline. Psychological science, 28(8), 1171-1179.

Salthouse, T. A. (2010). Selective review of cognitive aging. Journal of the International neuropsychological Society, 16(5), 754-760.

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. Bilingualism: Language and Cognition, 13(2), 231-252.

Schmid, M. S., & Jarvis, S. (2014). Lexical access and lexical diversity in first language attrition. Bilingualism: Language and Cognition, 17(4), 729–748. https://doi.org/10.1017/S1366728913000771

Schmid, M. S., & Keijzer, M. (2009). First language attrition and reversion among older migrants. International Journal of the Sociology of Language, 2009(200), 83–101. https://doi.org/10.1515/IJSL.2009.046

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. Linguistic approaches to bilingualism, 1(1), 1-33.

Sorace, A. (2016). Referring expressions and executive functions in bilingualism. Linguistic Approaches to Bilingualism, 6(5), 669 684.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. Psychological science, 10(3), 281-284.

Strijkers, K., Baus, C., Runnqvist, E., FitzPatrick, I., & Costa, A. (2013). The temporal dynamics of first versus second language production. Brain and language, 127(1), 6-11.

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. Cerebral cortex, 20(4), 912-928.

Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. Journal of memory and language, 65(4), 345-362.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. Proceedings of the National Academy of Sciences, 104(30), 12530–12535. https://doi.org/10.1073/pnas.0609927104

Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. International Journal of Bilingualism, 8(3), 257–277. https://doi.org/10.1177/13670069040080030601

Valente, A., Bürki, A., & Laganaro, M. (2014). ERP correlates of word production predictors in picture naming: a trial by trial multiple regression analysis from stimulus onset to response. Frontiers in neuroscience, 8, 390.

Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. Psychology and aging, 18(2), 332.

Wodniecka, Z., Szewczyk, J., Kałamała, P., Mandera, P., & Durlik, J. (2020). When a second language hits a native language. What ERPs (do and do not) tell us about language retrieval difficulty in bilingual language production. Neuropsychologia, 141, 107390.

Wolna, A., Łuniewska, M., Haman, E., & Wodniecka, Z. (2023). Polish norms for a set of colored drawings of 168 objects and 146 actions with predictors of naming performance. Behavior Research Methods, 55(5), 2706-2732.

Yagmur, K., Bot, K. D., & Korzilius, H. (1999). Language Attrition, Language Shift and Ethnolinguistic Vitality of Turkish in Australia. Journal of Multilingual and Multicultural Development, 20(1), 51–69. https://doi.org/10.1080/01434639908666369

Yilmaz, G., & Schmid, M. S. (2012). L1 accessibility among Turkish-Dutch bilinguals. The Mental Lexicon, 7(3), 249–274. https://doi.org/10.1075/ml.7.3.01yil__

Zhang, H., Diaz, M. T., Guo, T., & Kroll, J. F. (2021). Language immersion and language training: Two paths to enhanced language regulation and cognitive control. Brain and Language, 223, 105043.

8 . Appendices

A. Exploratory analysis

To check if the effect of L1 reimmersion was limited to the positivity, investigated in the a priori analysis, we conducted an exploratory analysis. We divided the data into five time windows; 0–60 ms, 60-140 ms, 140–250 ms, 250-350 ms, and 350–500 ms. Additionally, we used ten electrode cluster: Frontal (Fp1, AF3, AF4, Fp2), LeftAnterior (F7, F3, FC5), Fronto-central (Fz, FC1, FC2), RightAnterior (F8, F4, FC6), Left-central (C3, CP5), Centro-Parietal (Cz, CP1, CP2, Pz), Right-central (C4, CP6), Left-parietal (P7, P3, PO3), Right-parietal (P8, P4, PO4), Occipital (O1, Oz, O2).

A 2 × 2 within-subjects ANOVA was conducted considering Group and Context in all time windows and scalp areas.

Table A.1 shows the significance of the Group:Context interaction for each time window and electrode cluster. The effect seems to arise already during the earliest time window and is more pronounced on electrodes on the left hemisphere. Overall though the Group:Context interaction still have a broad spread over most scalp areas and time windows. Therefore, it seems that the change in the language environment can be traced through many processes involved during word production, from early processing steps to later ones around 350ms.

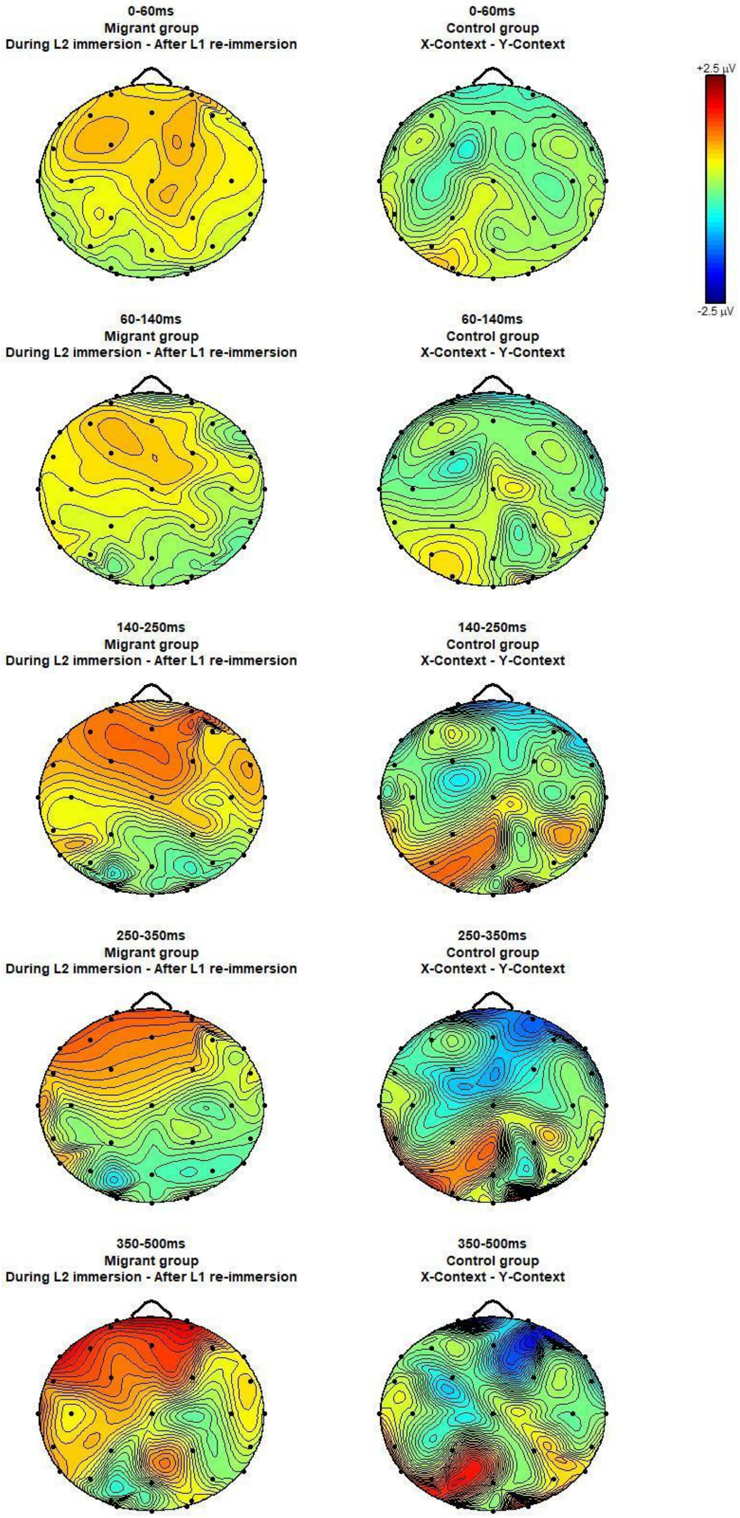
Table A.1

Statistics resulting from the ANOVAs conducted at each temporal window and scalp area for the Group:Context interaction.

Electrode cluster	0-60	60-140	140-250	250-350	350-500
Frontal	$F_{(1,1)} = 3.53$, $p = 0.07$	$F_{(1,1)} = 0.94$, $p = 0.34$	$F_{(1,1)} = 1.17$, $p = 0.28$	$F_{(1,1)} = 2.33$, $p = 0.13$	$F_{(1,1)} = 2.10$, $p = 0.15$
LeftAnterior	$\mathbf{F_{(1,1)} = 6.20}$, $\mathbf{p = 0.02}$	$\mathbf{F_{(1,1)} = 5.32}$, $\mathbf{p = 0.03}$	$\mathbf{F_{(1,1)} = 5.99}$, $\mathbf{p = 0.02}$	$\mathbf{F_{(1,1)} = 7.03}$, $\mathbf{p = 0.01}$	$\mathbf{F_{(1,1)} = 4.90}$, $\mathbf{p = 0.03}$

Fronto-central	$F_{(1,1)} = 3.73$, $p = 0.06$	**$F_{(1,1)} = 5.49$, $p = 0.02$**	**$F_{(1,1)} = 6.43$, $p = 0.01$**	**$F_{(1,1)} = 4.50$, $p = 0.04$**	$F_{(1,1)} = 1.36$, $p = 0.25$
RightAnterior	$F_{(1,1)} = 1.60$, $p = 0.22$	$F_{(1,1)} = 0.39$, $p = 0.53$	$F_{(1,1)} = 2.08$, $p = 0.16$	$F_{(1,1)} = 1.09$, $p = 0.30$	$F_{(1,1)} = 1.49$, $p = 0.23$
Left-central	**$F_{(1,1)} = 6.83$, $p = 0.01$**	**$F_{(1,1)} = 5.48$, $p = 0.02$**	**$F_{(1,1)} = 5.99$, $p = 0.02$**	$F_{(1,1)} = 3.91$, $p = 0.05$	$F_{(1,1)} = 2.23$, $p = 0.14$
Centro-Parietal	**$F_{(1,1)} = 8.17$, $p = 0.01$**	**$F_{(1,1)} = 5.52$, $p = 0.02$**	**$F_{(1,1)} = 4.75$, $p = 0.03$**	$F_{(1,1)} = 1.86$, $p = 0.18$	$F_{(1,1)} = 1.33$, $p = 0.25$
Right-central	**$F_{(1,1)} = 5.24$, $p = 0.03$**	$F_{(1,1)} = 2.46$, $p = 0.12$	**$F_{(1,1)} = 6.28$, $p = 0.02$**	$F_{(1,1)} = 1.31$, $p = 0.26$	$F_{(1,1)} = 1.03$, $p = 0.31$
Left-parietal	**$F_{(1,1)} = 13.98$, $p < 0.001$**	**$F_{(1,1)} = 6.18$, $p = 0.02$**	**$F_{(1,1)} = 4.12$, $p = 0.05$**	$F_{(1,1)} = 2.68$, $p = 0.11$	$F_{(1,1)} = 2.41$, $p = 0.13$
Right-parietal	**$F_{(1,1)} = 8.90$, $p < 0.01$**	$F_{(1,1)} = 1.35$, $p = 0.25$	$F_{(1,1)} = 1.46$, $p = 0.23$	$F_{(1,1)} = 0.88$, $p = 0.35$	$F_{(1,1)} = 1.48$, $p = 0.23$
Occipital	**$F_{(1,1)} = 10.86$, $p < 0.01$**	$F_{(1,1)} = 3.92$, $p = 0.05$	$F_{(1,1)} = 3.13$, $p = 0.08$	$F_{(1,1)} = 2.77$, $p = 0.10$	$F_{(1,1)} = 1.32$, $p = 0.26$

Figure A.1



Note. Scalp topography maps for all time windows of the exploratory analysis for the migrant group (left column) and the control group (right column).

B. Gender analysis

To ensure that our analysis was not affected by the disproportionate gender distribution in our participant sample, we analysed the ERP data involving gender as a fixed factor. Its inclusion did not significantly change our results with the P2 model (Table B.1) retaining its significant interaction between Group and Context. In the same line, there were no changes for the N300 model (Table B.2). Altogether, this indicates that gender did not play a significant role in how the language system adapts to changes in the language environment. However, given the uneven gender distribution in our sample, more studies involving a more balanced participant group are still needed.

Table B.1

Fixed effects for the LME model with ERP amplitude and both groups for the P2 component including Gender as a fixed factor.

	Estimate	SE	t	by-Picture SD	by-Participant SD
Intercept	-0.06	0.08	-0.69	0.10	0.31
Group	-0.09	0.11	-0.87		-
Context	0.03	0.03	0.94		0.16
Word-lexical frequency	0.01	0.04	0.24	-	
Participant's Gender	-0.11	0.11	-1.06		-

Participant’s Age	0.23	0.05	4.84***		-
Age of L2 acquisition	-0.08	0.05	-1.83		-
L2 proficiency	-0.06	0.04	-1.34		-
log (Trial number)	0.01	0.02	0.40		
Group: Context	0.17	0.06	2.73**		-
Group: Word-lexical frequency	0.13	0.08	1.57	-	-
Control Group : AfterL1-Immersion : Word-lexical frequency : Female	0.05	0.05	0.88	-	-
Migrant Group : AfterL1-Immersion : Word-lexical frequency : Female	-0.09	0.08	-1.24	-	-
Control Group : DuringL2-Immersion : Word-lexical frequency : Female	0.09	0.05	1.78	-	-
Experimental Group : DuringL2-Immersion : Word-lexical frequency : Female	-0.06	0.08	-0.84	-	-
Control Group : AfterL1-Immersion : Word-lexical frequency : Male	0.08	0.06	1.41	-	-
Experimental Group : AfterL1-Immersion : Word-lexical frequency : Male	-0.06	0.10	-0.54	-	-

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.2

Fixed effects for the LME model with ERP amplitude and both groups for the N300 component including Gender as a fixed factor.

	Estimate	SE	t	by-Picture SD	by- Participant SD
Intercept	-0.08	0.09	-0.89	0.23	0.39
Group	-0.20	0.13	-1.51		-
Context	0.02	0.04	0.69	0.04	0.20
Word-lexical frequency	0.03	0.04	0.70	-	
Participant's Gender	-0.05	0.13	-0.40		-
Participant's Age	0.07	0.06	1.17		-
Age of L2 acquisition	-0.07	0.06	-1.21		-
L2 proficiency	0.01	0.05	0.27		-
log (Trial number)	0.02	0.02	1.15		
Group: Context	0.11	0.07	1.54		-
Group: Word-lexical frequency	0.06	0.08	0.69	-	-
Control Group : AfterL1-Immersion : Word-lexical frequency : Female	0.01	0.05	0.10	-	-
Migrant Group : AfterL1-Immersion : Word-lexical frequency : Female	-0.06	0.07	-0.80	-	-

Control Group : DuringL2-Immersion : Word-lexical frequency : Female	0.03	0.05	0.57	-	-
Experimental Group : DuringL2-Immersion : Word-lexical frequency : Female	-0.03	0.07	-0.35	-	-
Control Group : AfterL1-Immersion : Word-lexical frequency : Male	0.04	0.06	0.80	-	-
Experimental Group : AfterL1-Immersion : Word-lexical frequency : Male	0.05	0.10	0.50	-	-

Investigation of long- and short-term adaptations of the bilingual language system to different language environments: evidence from the ERPs

Alba Casado 1,2*, Jonas Walther 1,3*, Agata Wolna 1,4, Jakub Szewczyk 1, Antonella Sorace 5 & Zofia Wodniecka 1

1. Psychology of Language and Bilingualism Lab, Institute of Psychology, Jagiellonian University, Krakow, Poland

2. Mind, Brain, and Behavior Research Center (CIMCYC); Department of Experimental Psychology; University of Granada, Granada, Spain

3. Department of Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany

4. McGovern Institute for Brain Research, MIT, Cambridge, MA, 02139, USA

5. School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, United Kingdom

*Joint first authorship

Address for correspondence

Alba Casado, e-mail – albacasado@ugr.es

Mind, Brain, and Behaviour Research Center (CIMCYC), Department of Experimental Psychology, University of Granada. Campus Cartuja. C/ Profesor Clavera s/n, 18011, Granada, Spain

Zofia Wodniecka, e-mail – zofia.wodniecka@uj.edu.pl

Psychology of Language and Bilingualism Lab, Institute of Psychology, Jagiellonian University in Krakow, Ul. Ingardena 6, 30-060

Style Definition ...
Style Definition ...
Style Definition ...
Style Definition ...
Style Definition ...
Style Definition ...
Style Definition ...
Style Definition ...
Style Definition ...
Style Definition ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...

Formatted: Header

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Double

Formatted: Footer

Abstract

Does a long-term stay in a foreign language country affect ~~our ability to retrieve words~~word retrieval in our native language? And if so, are the effects reversible? The present study explored the neural correlates of single~~-~~word production in the native language and their dynamics due to two types of changes in the language environment: long-term immersion in a foreign language (L2) environment and ~~a~~ short-term reimmersion in a native language (L1) environment. We tested Polish-English migrants living in the UK (L2 environment) for an average of ten years and Polish-English controls living in Poland (L1 environment). All participants performed ~~a~~an L1 picture-naming task ~~in their L1~~ while we recorded their electrophysiological responses. The migrants were tested before and after visiting the L1 environment, while the controls were tested twice in their L1 environment. Our focus was on two event-related components previously associated with the ease of lexical access: P2 and N300. We found no modulations related to N300, but some in the P2 time~~-~~ window, although their distribution was more frontal than previously reported. ~~Although there~~There was no main effect ~~related to~~of the long-term ~~migration~~immersion in the L2 environment, suggesting that the effectiveness of producing words in L1 was similar across the two groups. However, the short-term change in the language environment modulated the early positivity in migrants: smaller frontal positivity was reported in response to picture naming after the short reimmersion in the L1 environment than during the L2 immersion. These results indicate that the short-term changes in the language environment induce modulations in the neural response, which may reflect higher proactive control applied in L1 production during L2 immersion and its reduction after ~~a~~ short-term L1 immersion.

Formatted: Header

Formatted: Font: Bold, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Line spacing: Double

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

Keywords: bilingual lexical access; L2 immersion; L1 reimmersion; P2; N300.

Formatted: Header

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Double

Formatted: Footer

~~1.~~

1. **Introduction**

Moving to a foreign country with a language different from the native one usually requires the speaker to use the second language (L2) ~~on a~~ daily~~ basis~~. Being immersed in an environment where L2 is predominant can affect the ~~speakers'~~speakers' language system. While speakers immersed in an L2 environment have plenty of opportunities to hone their L2, they have fewer opportunities to use their native language (L1~~)~~), and there is a reduced variability of L1 interlocutors. These two factors are likely responsible for the fact that speakers immersed in L2 environments often experience difficulties producing words in L1 (Ammerlaan, 1996; Hulsen, 2000; Linck et al., 2009; Schmid & Jarvis, 2014; Schmid & Keijzer, 2009; Yagmur et al., 1999), especially low-frequency words (Baus et al., 2013; Botezatu et al., 2021). These difficulties using L1 are usually discussed in the literature as an overall decrease in L1 access or availability. Alternative explanations relate to more~~-~~ fundamental changes in the lexicon or general language knowledge due to L2 immersion. Overall, the phenomenon has been ~~referred to as 'language attrition' (Köpke & Schmid, 2004; Schmid, 2009).~~called 'language attrition' (Köpke & Schmid, 2004; Schmid, 2009). Still, the decrease in L1 accessibility due to L2 immersion is not experienced in all linguistic domains such as speech fluency or language comprehension in morphosyntactic processing (Bergmann et al., 2015; Gnitiev & Bátyi, 2022; Schmid, 2009). The decrease in L1 accessibility also does not seem to be permanent, as previous studies showed that reimmersion in the native language environment ~~can help regain levels~~leads to a return of language skills ~~from~~to a level before L2 immersion (Chamorro et al., 2016; Köpke & Genevska-Hanke, 2018; Linck et al., 2009). However, not all studies report a behavioural effect of long-term L2 immersion on L1 access (Beatty-Martínez et al., 2020; Casado, Walther, et al., 2023; Yilmaz & Schmid, 2012). It is

Formatted: Header

Formatted: Font: 12 pt, Bold, Font color: Text 1

Formatted: Normal, Justified, Indent: Hanging: 0.19", Space Before: 12 pt, After: 12 pt, Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1, Not Highlight

Formatted: Font color: Text 1, Not Highlight

Formatted: Footer

Formatted: Header

therefore of interest to explore whether the absence of behavioural effects can still be accompanied by signs of difficulty on the neural level.

Given that none of the previous studies have employed electrophysiological indices to track changes due to language immersion in language processing at the brain level, our goal was to fill-in the gap in the literature. The electrophysiological (EEG) technique is a tool ~~allowing~~that allows one to explore brain reactions to stimuli with ~~a very precise~~high time resolution. As a result, event-related potentials (ERPs) can provide a measure of the processing of stimuli even when there is no behavioural change ~~(Bice & Kroll, 2015; McLaughlin et al., 2004). It is, therefore, of interest~~(Bice & Kroll, 2015; McLaughlin et al., 2004). Thus, EEG allows us to ~~explore~~investigate whether the difficulty in accessing the native language in migrants is reflected by ~~electrophysiological~~the ERP markers and whether the same markers are also sensitive to short-term reimmersion to the L1. Consequently, in the current study, we examined the electrophysiological response in the picture naming of Polish-English migrants in an L1 (Polish) and an L2 (English) environment. Below, we briefly review the differing results of studies on how L1 lexical access can be affected by the language environment.

Formatted: Font color: Text 1, Not Highlight

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1, Not Highlight

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1, Not Highlight

Formatted: Font color: Text 1, Not Highlight

Formatted: Font color: Text 1, Not Highlight

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

1.1. Previous studies exploring long-term bilinguals' L1 lexical access in different language environments

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Justified, Space Before: 14.95 pt, After: 14.95 pt

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Font: 12 pt, Font color: Text 1

The data on the effect of L2 immersion are somewhat inconclusive. On the one hand, previous studies reported no differences in the speed of lexical access to L1 when L1 access is compared between groups of bilinguals immersed in L1 and bilinguals immersed in L2 for a long time, at least when measured using a relatively simple task such as picture naming (Beatty-Martinez et al., 2020; Casado, Walther, et al., 2023; Yilmaz and Schmid, 2012). On the other hand, when

Formatted: Footer

lexical L1 access is compared in the same L2-immersed individuals before and after reimmersion to an L1 environment, there seems to be a difference in lexical L1 access; in particular, after ~~a~~ short reimmersion in an L1 environment the L1 lexical access is easier (Baus et al., 2013; Botezatu et al., 2021; Linck et al., 2009). Our ~~own~~ behavioural data (Casado, Walther, et al., 2023) demonstrated that when the speed of picture naming in L1 was assessed, Polish-English bilinguals immersed in the L2 environment (United Kingdom) for about 10 years (2 - 24 years) performed similarly as Polish-English bilinguals living in the L1 environment (Poland). Altogether, based on the published studies, as far as picture naming speed is concerned, bilinguals who reside in the L2 environment for a relatively long period ~~of time~~, appear to have L1 words equally available as bilinguals who reside in the L1 environment.

~~Despite the fact that~~Even though no clear detrimental effect of long-term L2 immersion for L1 lexical access has been observed in comparison with speakers in the L1 environment, some studies reported that when L2~~-~~ immersed speakers are reimmersed in their L1 environment, i.e., revisit their home country for a short time, their speech performance in L1 is facilitated, compared to the time before the visit. This has been shown by faster naming latencies during a blocked picture-naming task in L1 (Casado, Walther, et al., 2023). In particular, in the latter study, we found that Polish-English bilinguals living in the UK showed faster naming latencies in high-frequency words after ~~a~~ short reimmersion in Poland, the L1 environment, than during immersion in the UK, the L2 environment. We argued that during their reimmersion in the L1 environment, the speakers primarily encountered high-frequency words, which led to the facilitation of lexical access to these words~~, an~~. This effect ~~that~~ was observed even a few days after their return to the L2 environment.

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: Not Italic, Font color: Text 1

Formatted: Font color: Text 1, Not Highlight

Formatted: Font color: Text 1, Not Highlight

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: Not Italic, Font color: Text 1

Formatted: Footer

Altogether, based on the behavioural evidence, we can distinguish two ~~different~~distinct aspects of L1 access in migrant bilinguals. On the one hand, it seems that long-term migrants do not necessarily experience difficulties in L1 lexical access compared to bilinguals in their L1 environment (Beatty-Martinez et al., 2020; Casado, Walther, et al., 2023; Yilmaz & Schmid, 2012). On the other hand, the frequency of encountering L1 words or using L1 structures in each language environment impacts the efficiency of accessing L1. Thus, even though long-term immersion in the L2 environment does not necessarily hamper L1 access (at least when assessed via picture naming latencies), migrants show clear sensitivity to changes in language environment: ~~a~~ short reimmersion to the native environment appears to temporarily facilitate retrieval of words in L1, especially those that are more frequently used. Still of note is that the assessment of L2 and L1 immersion above was conducted as a between-group comparison of migrants and residents in L1. At the same time, the effects of L1 reimmersion were assessed as a within-group comparison in the migrant population. The conflicting results could therefore result from individual differences and a lack of sensitivity of the used methodology. Considering that previous research exploring L1 access in L2~~-~~-immersed bilinguals involved uniquely behavioural measures, it is unknown whether we could detect more subtle differences on the neural level, specifically when comparing L2~~-~~-immersed bilinguals with bilinguals in the L1 environment. As stated above, ERP measures might provide ~~an~~ increased sensitivity ~~in comparison to~~than behavioural measures even when there is no behavioural change (Bice & Kroll, 2015; McLaughlin et al., 2004) and give us additional ~~insight~~insights into the processes underlying bilingual language production.

1.2. Electrophysiological correlates of lexical access

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Justified, Space Before: 14.95 pt, After: 14.95 pt

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Footer

Formatted: Header

Previous studies identified two ~~ERPs~~ERP components that have shown sensitivity to word retrieval difficulty in the picture-naming task.

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

The first component is the production P2, characterised by a positive activity peaking around 200 ms in central~~[1] electrodes (Baus et al., 2020; Strijkers et al., 2010, 2011, 2013). Larger~~ electrodes (Baus et al., 2020; Strijkers et al., 2010, 2011, 2013). More positive P2 amplitudes have been related to more difficult lexical access. For instance, low-frequency words in L1, which take longer to retrieve than high-frequency words in L1, have been found to evoke ~~higher~~more positive amplitudes of the P2 component than high-frequency words (Strijkers et al., 2010). Along the same lines, retrieval of words in L1 with a high level of interlexical competition (evoked by cumulative semantic interference) has been related to ~~higher~~more positive amplitudes of the P2 component (Costa et al., 2009). Additionally, L2 words, which in unbalanced bilinguals take longer to retrieve than L1 words, were associated with ~~higher~~more positive amplitudes of the P2 component (Strijkers et al., 2013). Altogether, previous studies ~~showed~~indicate that the naming P2 component ~~reflects~~may reflect difficulties in lexical access during word production at the lemma selection stage, according to Indefrey and Levelt's model (2004) of language production.

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Another component previously associated with the difficulty of retrieving words during picture naming is the N300. It is characterised by a negative activity peaking around 300 ms in central electrodes. Previously the component has been associated with the recognition of pictures showing

---

[1] Previous studies characterised the P2 component at different locations depending on the reference electrode(s) chosen in the preprocessing of the signal: parietal-occipital (Costa et al., 2009, nose reference), posterior with a wide distribution (Strijkers et al., 2010, nose reference), and fronto-central (Strijkers et al., 2011, left mastoid reference). When re-referencing the signal to the mean of mastoid electrodes, Strijkers et al. (2013) found a posterior distribution for word-lexical frequency effect and a broadly central distribution for the L1-L2 comparison effect. Therefore, we believe that the exact location of the P2 component might depend on the exact experimental manipulation and the used reference electrode. Given the similarity of the design, we decided to follow Strijkers et al (2013) reference – the mean of the two mastoid electrodes– in the preprocessing of the signal and to select their electrode distribution of the P2 component as a template.

Formatted: Footer

~~smaller~~less negative amplitudes for non-ambiguous pictures and repeated pictures (Curran et al., 2002; Federmeier & Kutas, 2001; Henson et al., 2004; Philiastides et al., 2006; Schendan & Kutas, 2002, 2007; van Petten et al., 2000). The component is also sensitive to the difficulty of integrating concepts with pictures (Barrett & Rugg, 1990; Eddy et al., 2006; Gratton et al., 2009; Holcomb & Mcpherson, 1994; Philiastides & Sajda, 2006; Voss et al., 2010; West & Holcomb, 2002). Recently, the N300 component was identified as reflecting effort associated with the difficulty of accessing the lexical representation of the to-be-named picture, the task at hand (Wodniecka et al., 2020). More specifically, the ~~increased~~more negative N300 ~~amplitude was~~amplitudes were observed for naming pictures in more difficult task conditions, e.g. when pictures were named for the first time in the experiment, compared to when they were repeated, or when pictures were named in L1 after a block of previous naming other pictures in L2. That is, the ~~higher~~more negative N300 amplitudes seem to reflect difficulties ~~retrieving~~integrating the lexical information associated with the names of corresponding pictures~~.~~ with the perceptual features of the image.

Altogether, we speculate that during picture naming two different components could reflect the difficulty of accessing the picture's name at a different timing; on the one hand, there is the naming P2 around the 200 ms, associated with the lemma selection stage which could be modulated by the amount of competing information active by the lexical concept (Costa et al., 2009; Strijkers et al., 2010; 2013). On the other hand, there is the N300 around 300 ms, associated with the difficulty of integrating the process of picture recognition in synchrony with the word retrieval process of the corresponding name in the target language (Wodniecka et al., 2020). Thus, the N300 component underlies the integration of two subprocesses: visual features that allow the recognition of the picture and the task at hand, which in this case is accessing the picture's name. Therefore, accessing the lexical information (which takes place in the previous time window) is needed to obtain

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

modulations in the following one (N300) corresponding to the integration of the visual and the lexical information (see Valente et al., 2014).

1.3. Current study

In the current study, we explored electrophysiological correlates of lexical access difficulty in ~~the~~ native language production associated with (1) long-term immersion in L2; and (2) the effects of ~~a~~ short-term reimmersion in the L1. For this aim, we recorded EEG data during a picture naming task in L1 (Polish). The task was part of a large-scale study in which we tested a migrant group of Polish-English bilinguals living in the UK (L2 environment) and a control group of Polish-English bilinguals ~~living~~residing in Poland (L1 environment). Each group was tested twice with a mean between-session interval of 94 days (*SD* = 48, 30 - 259 days). The migrant group was tested once during immersion in the L2 environment and once after ~~a~~ short reimmersion in the L1 environment with a variable interval between sessions. The control group was tested both times in the L1 environment, with a similar interval between the tests as the migrant group. The detailed pattern of behavioural data from the picture naming task was already reported in Casado, Walther, et al. (2023). Here we focus solely on the electrophysiological indices of picture naming performance.

We built on previous studies that identified two ERP components during picture naming ~~which were~~ related to the difficulty of word retrieval during language production: the naming P2 component and the N300 component, both indicators of lexical access difficulty. Based on these results, we expected to observe these components to reflect changes in ~~migrants’ cognitive~~ migrants' language system, also in comparison with control bilinguals, due to changes in the language environment.

Formatted: Header

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font: Not Italic, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: Not Italic, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

We made three predictions. Firstly, given that the ERPs can reflect neural modulations even in the absence of any behavioural effects, we predicted that naming pictures in L1 ~~will~~would be accompanied by an enhanced amplitude of P2 and N300 in the migrant group compared to the control group. Such a result would suggest that the difficulty that migrants experience while retrieving words in L1 is reflected in brain activity during lexical access and its integration with picture recognition~~ and naming~~. Our second prediction concerned the effects of short-term L1 reimmersion. We predicted reduced P2 and N300 amplitudes after L1 reimmersion compared to the L2 immersion condition. The third prediction related to the interaction between language environment and frequency of words to be named: we expected that the P2 amplitudes should also be modulated by the lexical frequency of the words; in particular, low-frequency words should evoke ~~higher~~more positive amplitudes of the P2 compared with high-frequency words. Finally, a potential differentiation in the effects of P2 and N300 regarding the difference between groups and between before and after L1 reimmersion would give us insight into the time~~-~~ course of the ~~effect~~phenomenon during word production. Modulations in the P2 amplitude would indicate changes to the lexical processing steps during word production caused by different language activation levels. On the other hand, modulations in the N300 amplitude could be more closely attributed to the integration of visual and the already active lexical information during picture naming. Therefore, a modulation in the P2 but not in the N300 for the migrant group would indicate that migrants showed greater difficulty accessing L1 ~~words’~~words' lexical information, and not a difficulty in integrating visual information during L1 word production, reflected in the later component.

2. ~~2.~~Methods

2.1. Participants

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: List Paragraph, Justified, Space Before: 12 pt, After: 12 pt, Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

Formatted: Font: 12 pt, Bold, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Justified, Space Before: 14.95 pt, After: 14.95 pt

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Footer

We analysed data collected from a ~~relatively~~rather large-scale study in which we tested the consequences of immersion and reimmersion ~~in~~into many language domains by collecting the data from a battery of language-related tasks. In that study, we collected data from various tasks from ~~a sample of~~ 55 Polish-English bilinguals living in the United Kingdom (migrant group) and 56 Polish-English bilinguals residing in Poland (control group). Each group was tested twice with a mean between-session interval of 94 days (*SD* = 48, 30 - 259 days). In the present study, we focus on L1 lexical access, assessed through the EEG response to naming pictures in L1, which was available for 32 migrant bilinguals (see details below) and contrasted with a matched sample of 32 control bilinguals.

For the migrant group, we recruited Polish native speakers living in the UK for ~~a minimum of~~at least two years. As indicated before, the migrant group was tested twice: ~~“~~"During L2 immersion~~”,~~", that is after at least 30 days fully immersed in the L2 environment, without leaving the UK; and ~~“~~"After L1 reimmersion~~”,~~", after a reimmersion in the L1 environment (Poland), less than 7 days after returning from their L1 environment. From the initial sample of 55, we excluded nine participants who did not complete the session during L2 immersion and another five because they did travel to Poland within 30 days beforehand and did not fulfil the selection criteria. Additionally, nine more participants were excluded due to technical problems while recording the electrophysiological and/or behavioural responses. The final sample included in the analyses consisted of 32 participants[2].

For the control group, we recruited native speakers of Polish with ~~high~~upper-intermediate English proficiency. Moreover, pre-selection criteria only allowed those participants who spent ~~the last~~ 30

Formatted: Header

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

---

[2] All the migrant group participants included in the present analysis were also included in the behavioural analysis presented in Casado et al. (2023).

days ~~in Poland~~ before each session in Poland. Like the migrant group, the control group completed two sessions, the ~~“~~"X-Context~~”~~" session and the ~~“~~"Y-Context~~”~~" session. In both sessions, they were immersed in the L1 environment, that is, tested in Poland. By ~~also~~ having two sessions for the control group as well, we ~~were able to~~could account for the effects of repetition of the picture naming task within a balanced design. Out of the initial group of 56, seven participants were excluded because they completed only one experimental session, seven additional participants were excluded due to ~~the~~ bad quality of the EEG recordings, and three more participants were excluded due to recording errors. The remaining 32 participants[3] were compared to the migrant group.

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

---

[3] Only 25 out of 32 control group participants were analysed in both Casado, Walther, et al. (2023) and the present study, given that only these data sets had both high-quality behavioural and electrophysiological data.

Formatted: Footer

Table 1

Demographic information and language experience of participants.

	Migrant group (N = 32)		Control group (N = 32)		t-test	
N	32 (29 female)		32 (23 female)			
Participant's Age (years)	36.16~~ (6.45)~~ [33.83, 38.48]		29.84~~ (7.35)~~ [27.13, 32.56]		t(62) = -4.19, p = <0.01***	
SES (1 - 10)	6.64~~ (1.54)~~ [6.09, 7.20]		5.91 [5.33, 6.~~00 (1.65)~~48]		t(62) = -1.62, p = 0.11	
Years of education	18.35~~ (2.58)~~ [17.43, 19.28]		17.~~00 (1.98)~~16 [16.39, 17.92]		t(62) = -2.34, p = 0.02*	
Length of residence in L2 environment (years)	9.66~~ (4.86)~~ [7.90, 11.41]		-			
Length of reimmersion in L1 environment (days)	13.~~37 (8.18)~~38 [10.42, 16.33]		-			
Time delay between L1 reimmersion and recording (days)	3.06~~ (1.86)~~ [2.39, 3.74]		-			
Self-assessed language experience (1-10)	L1	L2	L1	L2	L1	L2
Self-rated proficiency	9.82~~ (0.47)~~ [9.65, 9.99]	7.86~~ (0.93)~~ [7.52, 8.19]	9.~~7343 (0.70)~~71 [9.46, 9.96]	7.~~20 (1.29)~~17 [6.74, 7.60]	t(62) = -0.57,	t(62) = -2.33,
					p = 0.57	p = 0.02*
Speaking	9.64~~ (0.79)~~ [9.36, 9.92]	7.68~~ (1.25)~~ [7.23, 8.13]	9.~~59 (1.07)~~56 [9.17, 9.95]	6.~~63 (1.52)~~59 [6.10, 7.09]	t(62) = -0.20,	t(62) = -3.02,
					p = 0.84	p = <0.01*

Formatted: Header

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Justified, Space Before: 12 pt, Line spacing: Double

Formatted: Font color: Text 1

Formatted: Justified, Line spacing: Double

Split Cells

Split Cells

Split Cells

Split Cells

Formatted: Footer

Formatted: Header

Writing	9.77 ~~(0.75)~~ [9.50, 10.04]	7.42 ~~(1.29)~~ [6.96, 7.89]	9.~~72 (0.63)~~66 [9.40, 9.91]	6.~~75 (1.76)~~66 [6.04, 7.27]	t(62) = -0.28,	t(62) = -1.74,
					p = 0.78	p = 0.09
Listening	9.94 ~~(0.25)~~ [9.85, 10.02]	7.91 ~~(0.86)~~ [7.60, 8.22]	9.84 ~~(0.51)~~ [9.66, 10.03]	7.53 ~~(1.29)~~ [7.12, 7.94]	t(62) = -0.91,	t(62) = -1.37,
					p = 0.37	p = 0.17
Reading	9.93 ~~(0.25)~~ [9.85, 10.02]	8.42 ~~(0.98)~~ [8.07, 8.77]	9.78 ~~(0.79)~~ [9.50, 10.07]	7.91 ~~(1.35)~~ [7.43, 8.39]	t(62) = -1.05,	t(62) = -1.74,
					p = 0.30	p = 0.09
Percentage of daily use (%)	40.46 ~~(15.94)~~ [34.72, 46.21]	59.25 ~~(15.~~ [53~~)~~.66, 64.85]	81.~~81 (16~~32 [75.44, 87.19~~)~~]	16.~~43 (11.13)~~84 [12.75, 20.93]	t(62) = 10.29,	t(62) = -12.68,
					p = <0.01***	p = <0.01***
Age of L2 acquisition (years)	-	13.05 ~~(3.72)~~ [11.70, 14.39]	-	~~9.38 (3.82)~~10.00 [8.41, 11.59]	t(62) = -3.89, p = <0.01***	
Intensity of language switching (1 = no switching, 10 = always switching)	4.82 ~~(2.55)~~ [3.90, 5.74]		4.~~31 (2.26)~~06 [3.27, 4.85]		t(62) = -0.84, p = 0.40	
Objective L2 proficiency measures						
LexTALE (mean accuracy in %)	-	~~77.82 (13.24)~~0.78 [0.73, 0.83]	-	~~73.12 (10.97)~~0.74 [0.70, 0.77]	t(62) = -1.55, p = 0.13	
General English Test (mean accuracy in %)	-	89.88 ~~(9.48)~~ [86.44, 93,28]	-	~~84.50 (7.31)~~85.12 [82.44, 87.84]	t(62) = -2.54, p = 0.01*	

Formatted: Footer

Note. The first part of the table describes the demographic information of the final migrant group and the final control group. The rows display (1) the number of participants with the number of women in brackets (gender was self-reported by participants), (2) age (in years), (3) socio-economic status on a 1 to 8 scale based on ~~Adler et al. (2000),~~Adler et al. (2000), (4) years of education (in years), (5) length of residence in an L2 environment (in years), (6) length of immersion the L1 environment and (7) time ~~delay~~passed between arrival at the ~~return from~~airport in the ~~L1 environment~~UK and the experimental recording session. The second part of the table summarises the self-assessed language experience based on a questionnaire. The self-rated proficiency is presented on a scale from 1 to 10, where 1 = "no knowledge of a given language" and 10 = "native-like proficiency". The daily use of each language is presented in percentages ~~and~~while the age of acquisition is presented in years. Bilingual switching is presented on a scale from 1 to 10, where 1= "I never switch languages within sentences" and 10 = "I always switch languages within sentences". The objective L2 proficiency measures in English are presented in percentages. ~~Standard deviations~~95% confidence intervals are given in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All participants received monetary compensation for participating in the study (65 pounds per session or the equivalent in Polish zloty) and a selection of Polish books. The study met the requirements and received the approval of the Ethics Committee of Jagiellonian University Institute of Psychology concerning experimental studies with human subjects.

All~~ the~~ participants learned English as a second language and used it ~~on a~~ daily~~ basis~~ (see Table 1), however, the experimental group used L2 significantly more than the control group. We assessed their English proficiency with the General English Test (by Cambridge Assessment: https://www.cambridgeenglish.org/test-your-english/general-english/) and an online version of the LexTALE task (Lemhöfer & Broersma, 2012) programmed in Inquisit (Inquisit 5 [Computer software] (2016). Retrieved from https://www.millisecond.com). The selection criteria for participating in the study were: a self-reported English proficiency of upper-mediate (B2) or above, 70% or higher accuracy in the General English Test, and 60% or higher accuracy in the LexTALE test. The migrant participants also reported that they only used Polish (never English or other languages) when contacting friends and family in Poland. Due to the small target population, we

Formatted: Header

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Justified, Line spacing: Multiple 1.15 li

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Not Italic, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

~~were~~could not ~~able to~~ recruit a balanced group of migrants ~~with respect to~~concerning gender, age, or language proficiency. Most likely due to the long-term immersion in the L2 environment, the migrant group showed slightly higher L2 proficiency than the control group, even though both of them only demonstrated on average an upper intermediate level of L2. The language experience of both groups is summarised in Table 1.

2.2. Task and procedure

The order of the sessions was counterbalanced for both groups of participants. In the final sample, the session during L2 immersion was the first session for 16 participants of the migrant group and the X-Context for 16 participants of the control group.

2.2.1. Materials

Table 2

Stimuli information for the four subsets used during picture naming.

	A	B	C	D
PL Age of Acquisition (years)	3.43 [3.27, 3.59]	3.56 [3.40, 3.73]	3.41 [3.28, 3.54]	3.49 [3.34, 3.64]
Name agreement (%)	0.93 [0.89, 0.96]	0.92 [0.88, 0.95]	0.95 [0.92, 0,97]	0.93 [0.90, 0.96]
Lexical frequency	2.77	2.71	2.72	2.85

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Justified, Space Before: 14.95 pt, After: 14.95 pt

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Justified, Space Before: 14.05 pt, After: 14.05 pt

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Justified, Space Before: 12 pt, Line spacing: Double

Formatted: Footer

	[2.59, 2.95]	[2.55, 2.88]	[2.54, 2.89]	[2.66, 3.04]
Phoneme length	5.58	5.38	5.72	5.60
	[5.20, 5.96]	[4.94, 5.82]	[5.32, 6.13]	[5.15, 6.05]

Note. The table shows the average stimuli properties of the used picture subsets. The rows display (1) the age of acquisition for the Polish picture names (in years) based on Haman et al. (2015), (2) the name agreement of the used pictures (in percent) based on Wolna et al. (2022), (3) lexical frequency based on Mandera et al. (2015), (4) phoneme length and (5) Imageability based on Wolna et al. (2022). 95% confidence intervals are given in parentheses.

The stimuli included in the picture-naming task consisted of 216 coloured images from the Cross-Linguistic Lexical Tasks database (Haman et al., 2017). We divided all ~~the~~ pictures into four subsets. ~~The~~These subsets of pictures were balanced for name agreement (based on Wolna et al., 2022), lexical frequency (based on Mandera et al., 2015), age of acquisition (Haman et al., 2015), and mean length in phonemes. Polish-English cognates and homophones were excluded during our stimuli selection process. Moreover, each subset contained a comparable number of images from different semantic categories. The four subsets were counterbalanced across participants and sessions so that no pictures were repeated between sessions and each participant ~~was seeing~~saw two out of four picture subsets.

The variation in the lexical frequency of pictures' target names ~~of pictures~~ allowed us to explore possible interactions between the L1 vs. L2 environments and the lexical frequency.

2.2.2. Procedure

Formatted: Header

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Multiple 1.15 li

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Justified, Space Before: 14.05 pt, After: 14.05 pt

Formatted: Footer

In the picture-naming task, pictures were displayed on a computer screen using DMDX (Forster & Forster, 2003). The pictures were presented in the centre of the screen. Each trial was preceded by a black screen presented for 1000 ms, followed by a fixation cross in the screen's centre for 1000 ms. A picture was then shown in the centre of the screen until the participant responded or until the ~~timeout~~time to respond was ~~reached~~over (3000 ms). The participants were instructed to name pictures aloud in their native language as quickly and accurately as possible. Each session of picture naming had a total of 58 trials (4 practice trials and 54 regular trials). Overall, the picture-naming task lasted approximately 5 ~~min~~minutes.

2.2.3 EEG Recording

The EEG was recorded during the picture-naming task at 1024 Hz from 32 Ag/AgCl scalp electrodes positioned at the standard 10-20 locations, mounted in an elastic cap, using the Biosemi Active Two recording system. Electrodes were initially referenced online to the Common Mode Sense electrode located at the C1 electrode. The horizontal and vertical electrooculogram (EOG) was recorded bipolarly using electrodes placed below and above a participant's left eye and at the outer canthus of each eye, respectively. The EEG signal was offline filtered with a band-pass filter (0.1 – 25 Hz frequency range; low cutoff slope: 24 dB/oct; high cutoff slope: 12 dB/oct) and re-referenced offline to the mean of the left and right mastoids. The data was offline preprocessed using BrainVision Analyzer (Brain Products, Gilching, Germany), downsampled to 256 Hz, and baseline-corrected. We extracted segments of 900 ms (-100 to 800). Ocular artefacts were removed with Independent Component Analysis (ICA, (Delormeet al., 2007; Jung et al., 2000)) by calculating the ICA components based on 1 Hz high-pass filtered data and removing the artifactual components. On average 2.8 components were removed per participant with ~~at most~~a maximum

Formatted: Header

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Justified, Space Before: 14.05 pt, After: 14.05 pt

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Footer

of 6 components being removed. The corrected ICA matrix was~~ then~~ applied to the 0.1 Hz high-pass filtered data set. Segments containing artefacts were cleaned manually. ~~Afterwards~~Afterward, the data was exported to Matlab for further analysis using EEGlab (Delorme & Makes, 2004) and ERPlab (Lopez-Calderon & Luck, 2014). Mean amplitudes were calculated with specific time windows and electrode selections for each ERP component, as described below.

2.3. Analysis

Two time~~-~~-windows were selected for the EEG analysis based on previous studies ~~(Costa et al., 2009; Strijkers et al., 2010, 2011, 2013; Wodniecka et al., 2020):~~(Costa et al., 2009; Strijkers et al., 2010, 2011, 2013; Wodniecka et al., 2020); the naming P2 around 175 ms (160 – 240 ms) and the N300 around 300 ms (250 – 350 ms). For each time window, the mean amplitude was calculated with a different selection of electrodes following the literature: FC1, FC2, Fz, Cz, CP1, CP2 for the naming P2 (Strijkers et al., 2013), and FC1, FC2, C3, C4, Cz, CP1, CP2 for the N300 (Wodniecka et al., 2020).~~. Additionally, due to temporal differences caused by the recording setup in each different location (display lag in lab in Poland where the control group was tested vs. lab in Edinburgh where the migrant group was tested), we applied a latency correction of 27 ms to the migrant group, effectively choosing earlier time windows for the migrant group for P2 (113 – 233 ms) and N300 (223 – 323 ms).~~

~~In order to~~To be able to relate the electrophysiological results of the present study with the findings of the behavioural analysis already reported in Casado, Walther, et al. (2023), we included ~~in the analyses~~ all~~ the~~ variables we controlled for in the previous study ~~so as~~in the analyses to maintain the two analyses as similar as possible. That is, besides the main interest factors (Group, Context, and Word-lexical frequency), we included some of the ~~participants'~~participants' variables the

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Justified, Space Before: 14.95 pt, After: 14.95 pt

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

Formatted: Header

groups were not matched for: the ~~participants’~~participants' age, and the age of L2 acquisition. Moreover, given that there were differences in L2 proficiency between groups, we also included this variable as a covariate in the main model.

Formatted: Font color: Text 1

Responses with naming latencies below 650 ms were removed from the data to avoid artefacts in the signal evoked by the articulatory movements (9.95%). We also excluded trials with inaccurate responses (4.24%). In total, 14.19 % of the trials were excluded[4].

Formatted: Font color: Text 1

Table 3

Summary of the final trial count after artefact rejection during EEG processing and removal of inaccurate responses and too early articulator movements.

Formatted: Font color: Text 1

Formatted: Justified, Line spacing: Multiple 1.15 li

Formatted: Font: Times New Roman, 12 pt, Font color: Text 1

	During L2 immersion/X-Context	After L1 reimmersion/Y-Context
Migrant group	88.6% (1531 trials)	82.0% (1417 trials)
Control group	87.0% (1504 trials)	85.6% (1479 trials)

Note. The table gives the percentage and number (in brackets) of remaining trials in the final analysis for each Group and Context. The raw data set contained 1728 trials within each condition.

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Justified, Line spacing: Multiple 1.15 li

Formatted: Font: Times New Roman, Font color: Text 1

We used R Studio (R Development Core Team, 2020, Version 4.0.2) to fit linear mixed-effects models with the lme4 library (Bates et al., 2015, Version 1.1-23). The general models included the mean amplitude of P2 and N300 as the dependent variables and participants and pictures as crossed random effects. As fixed effects, we included Group (Control, Migrant), Context (During L2

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

---

[4] To ensure that no group or context was disproportionately affected by trial rejection, we also looked at the trial numbers after rejection in each group and context. Due to the removal of trials with naming latencies below 650 ms, the migrant group in L2 immersion lost 7.12%, the migrant group after L1 reimmersion 12.15%, the control group in the X-Context 9.43% and the control group in the Y-Context lost 11.11% of their trials. Due to removing the erroneous trials, the migrant group in L2 immersion lost 4.24%, the migrant group after L1 reimmersion 4.28%, the control group in the X-Context 3.53% and the control group in the Y-Context lost 3.30% of their trials.

Formatted: Footer

immersion/X-Context, After L1 reimmersion/Y-Context), Word-lexical frequency (based on Mandera et al., 2015), ~~participant’s~~participant's Age, Age of L2 acquisition and L2 proficiency (based on the mean scores of the LexTALE and Cambridge proficiency tasks), and the Trial number. We also included the interactions between Group, Context, and Word-lexical frequency. Before running the analyses, all categorical predictors were deviation~~-~~-coded using a sum contrast (Group~~: ~~: Control group = -0.5, Migrant group = 0.5; Context: During L2 immersion/X-Context = -0.5, After L1 reimmersion/Y-Context = 0.5; Session order: first session = -0.5, second session = 0.5). The continuous predictors of Word-lexical frequency, ~~participant’s~~participant's Age, Age of L2 acquisition, and L2 proficiency were centred and scaled. The Trial number was log-transformed. The maximal model also included a by-~~Picture~~picture random intercept and random slopes for Group, Context, ~~participant’s~~participant's Age, Age of L2 acquisition, and L2 proficiency. Additionally, we used a by-Participant random intercept with random slopes for Context and Word-frequency.

We fitted the maximal models first (Barr et al., 2013). If the model did not converge, we first removed correlations between random effects and in the next step, the random effects with the smallest unique variance, following the recommendation by Bates et al. (2018). Summaries and p-values of the models were calculated using the lmerTest package (Kuznetsova et al., 2017). Absolute t-values greater than two were considered significant. For pairwise comparisons, the emmeans and lsmeans packages were used (Lenth, 2016, 2021).

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

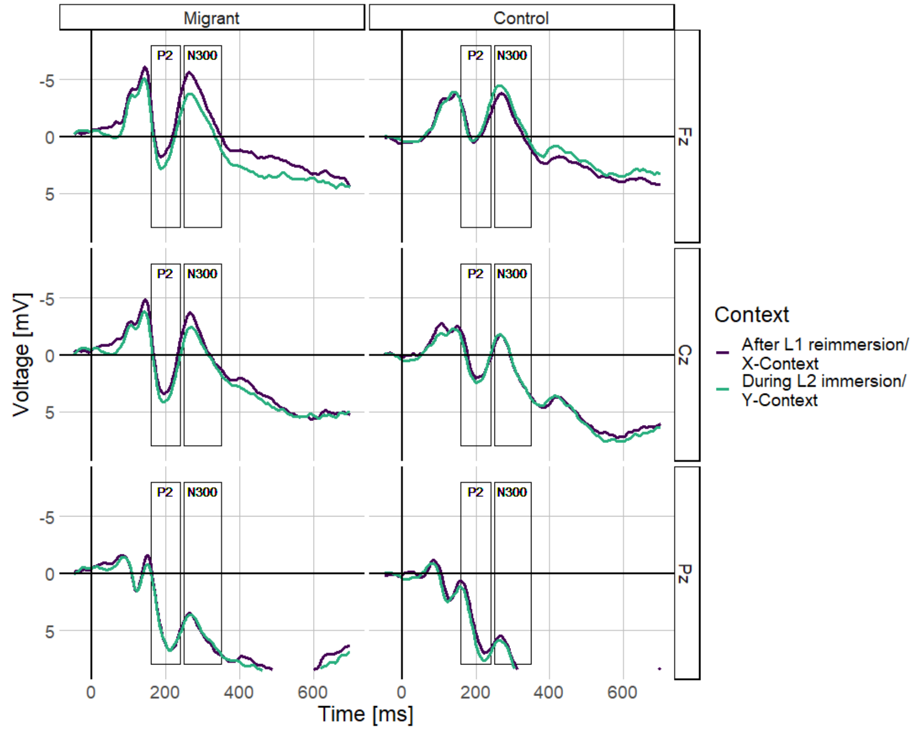
Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

3. Results

Figure 1



Formatted: Header

Formatted: Font: 12 pt, Bold, Font color: Text 1

Formatted: Font color: Text 1

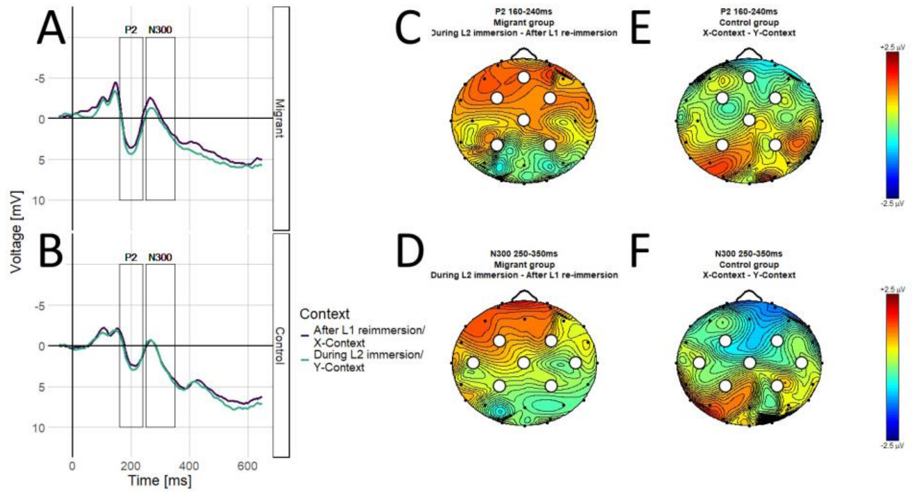
Formatted: Normal, Justified, Indent: Left: 0.5", Space Before: 12 pt, After: 12 pt, Line spacing: Double

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Double

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Footer

Formatted: Header



Note. Stimulus-locked grand average ERP waveforms for ~~naming P2 and N300 for~~ the Migrant **(A)** and Control ~~groups.~~**(B)** groups and scalp topography maps for naming P2 **(C, E)** and N300 **(D, F)**. Waveforms depict mean voltages ~~recorded from~~averaged over the ~~midline~~shared electrodes between the P2 and N300 clusters: FC1, FC2, Fz, Cz, CP1, and ~~Pz.~~

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

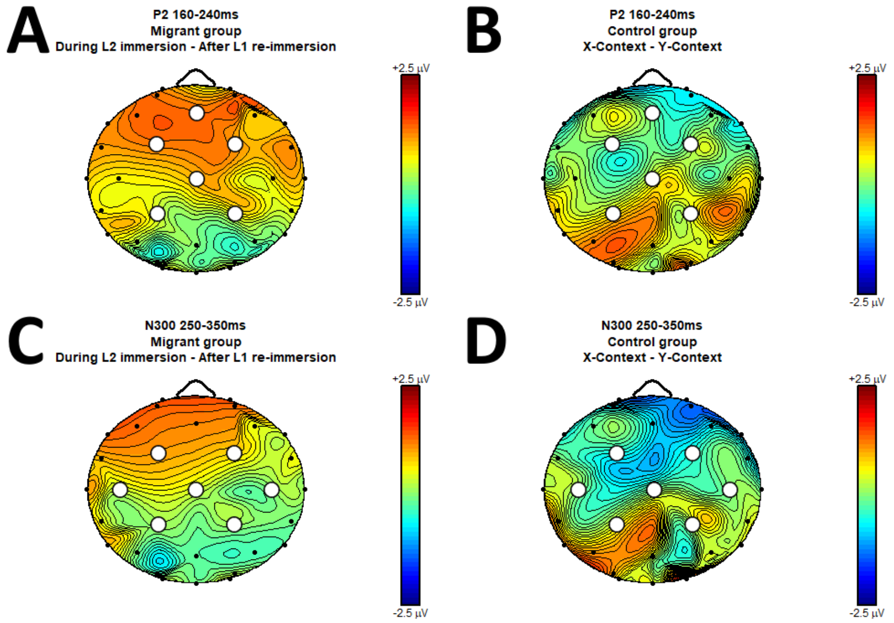
Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Footer

~~**Figure 2**~~



CP2. Scalp ~~topography maps for naming P2 **(A, B)** and N300 **(C, D)** and~~ plots show differences in voltage between different language contexts for the migrant group ~~**(A, C)**~~ and ~~the~~ between the two testing sessions for the control group ~~**(B, D).**~~ Selected electrode clusters used for calculating mean amplitudes are marked in white.

Formatted: Header

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Multiple 1.15 li

Formatted: Font: 11 pt, Not Italic, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Footer

The results are presented in Table ~~3~~4 and Table ~~4~~5. The detailed analysis of the behavioural data ~~were~~was reported in Casado, Walther, et al. (2023)[5]. Also). Additionally, a summary is presented in Table ~~5. For~~6. In the P2 component analysis, there was no main effect of ~~the~~ Group ~~but~~. Still, we found a significant interaction between Group and Context, demonstrating a modulation of the P2 component in the migrant group due to the language context: ~~higher~~more positive P2 amplitudes were observed during L2 immersion compared to after L1 reimmersion (see Figure 1 panel A). As expected, no differences in the P2 amplitudes between contexts were found for the control group (see Table ~~5~~6). However, it ~~needs to~~should be noted that the distribution of the P2 was more frontal than expected based on the previous studies (Costa et al., 2009; Strijkers et al., 2010; 2011; especially Strijkers et al., 2013).

Contrary to our predictions, we did not observe the effect of frequency on mean amplitude in the P2 component. Additionally, there was a significant effect of the ~~participant's~~participant's Age on the P2 amplitude with an increased amplitude for older participants.

In the analysis of the N300 component, we found no main effects of Group or Context (see Figure 1 panel B) and no significant interactions.

Table ~~3~~4

Formatted: Header

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Double

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Justified, Space Before: 12 pt, Line spacing: Double

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Footer

---

[5] The EEG analyses reported in the current paper were conducted on a subsample of participants reported in Casado, Walther, et al. (2023). To assure that the behavioural data pattern on the subsample converges with this reported in our previous paper, we re-ran the analyses on the subsample. The pattern was largely similar, with the main effect of the Group (beta = -0.06, t = -1.87, p = .07) and a Group and Context interaction approaching significance (beta = 0.04, t = 1.74, p = .09). However, the Group, Context and lexical frequency interaction was less pronounced (beta = 0.01, t = 0.89, p = .38). Of note, the removal of responses below 650 ms needed for EEG data analyses, reduced the number of trials analysed in the current analyses and might have impacted the absence of interaction with item frequency which we reported in Casado, Walther et al. 2023.

Fixed effects for the LME model with ERP amplitude and both groups for the P2 component.

	Estimate	SE	t	by- Picture SD	by- Participant SD
Intercept	-0.03	0.08	-0.44	0.10	0.31
Group	-0.05	0.10	-0.47		-
Context	0.03	0.03	0.86		0.16
Word-lexical frequency	0.00	0.01	0.22	-	
Participant’s Age	0.22	0.05	4.71***		-
Age of L2 acquisition	-0.09	0.05	-1.91		-
L2 proficiency	-0.07	0.04	-1.58		-
log (Trial number)	0.01	0.02	0.54		
Group: Context	0.18	0.06	2.76**		-
Group: Word-lexical frequency	0.00	0.02	-0.12	-	-
Control Group: Context: Word-lexical frequency	0.00	0.03	0.08	-	-
Migrant Group: Context: Word-lexical frequency	0.03	0.03	0.89	-	-

~~* *p* < 0.05, ** *p* < 0.01, *** *p* < 0.001~~

~~**Table 4**~~

Formatted: Header

Formatted: Font color: Text 1

Formatted: Justified, Line spacing: Double

Formatted Table

Formatted: Footer

Formatted: Header

Table 5

Fixed effects for the LME model with ERP amplitude and both groups for the N300 component.

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, Line spacing: Double

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Line spacing: Double

	Estimate	SE	t	by- Picture SD	by- Participant SD
Intercept	-0.09	0.09	-0.98	0.23	0.39
Group	-0.18	0.12	-1.50		-
Context	0.02	0.04	0.63		0.20
Word-lexical frequency	0.02	0.02	1.15	-	
Participant's Age	0.06	0.06	1.10		-
Age of L2 acquisition	-0.07	0.06	-1.26		-
L2 proficiency	0.01	0.05	0.21		-
log (Trial number)	0.03	0.02	1.35		
Group: Context	0.11	0.07	1.52		-
Group: Word-lexical frequency	0.00	0.02	-0.05	-	-
Control Group: Context: Word-lexical frequency	0.00	0.03	-0.10	-	-
Migrant Group: Context: Word-lexical frequency	0.02	0.03	0.61	-	-

Formatted: Footer

~~* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$~~

Table 6

~~**Table 5**~~

Summary of the electrophysiological data.

ERP component	Group	During L2 immersion / X-Context	After L1 reimmersion / Y-Context	Context comparison
P2	Migrant group	0.04 (0.07)	-0.08 (0.06)	*z* = -2.59; *p* = 0.01*
	Control group	-0.01 (0.07)	0.05 (0.06)	*z* = 1.33; *p* = 0.19
	Group comparison	*z* = -0.40; *p* = 0.69	*z* = 1.40; *p* = 0.16	
N300	Migrant group	-0.04 (0.08)	-0.11 (0.08)	*z* = -1.55; *p* = 0.12
	Control group	0.08 (0.08)	0.11 (0.08)	*z* = 0.62; *p* = 0.53
	Group comparison	*z* = 1.00; *p* = 0.32	*z* = 1.87; *p* = 0.06	

Note. The table gives the predicted amplitudes of the two ERP time windows for the two groups in the two Contexts and the pairwise comparison from their respective linear-mixed models ~~by~~ using the emmeans function. Standard errors are given in parentheses.

4. Discussion

In the present study, we explored neural signatures of L1 ~~lexical access in language~~ production ~~of~~by bilinguals in an L1 environment (native language country) and an L2 environment (foreign

Formatted: Header

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, Line spacing: Double

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Justified, Line spacing: Double

Formatted: Font: 11 pt

Formatted: Line spacing: Multiple 1.15 li

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Multiple 1.15 li

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Justified

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

language country). To this aim, we compared the EEG response evoked during an L1 picture-naming task of ~~a group of~~ Polish-English migrants immersed in the L2 environment with that of ~~a control group of~~ Polish-English bilinguals in the L1 environment. Moreover, we explored the ~~migrants’~~migrants' L1 lexical access under two different conditions: during L2 immersion, and after~~ a~~ short-term reimmersion in the L1 environment. We focused on two ERP components: P2 and N300, previously associated with the ease of lexical access during language production (Strijkers et al., 2010; 2011; Wodniecka et al., 2020).

We tested three predictions~~: first~~. First, if long-term immersion in the L2 environment results in decreased accessibility of L1 words, we should observe larger amplitudes of the P2 (more positive) and N300 (more negative) effects in the migrant group compared to the control group. Second, if ~~a~~ short-term reimmersion in L1 can reverse the decreased access to L1 words due to long-term immersion in L2, we should observe lower amplitudes of the P2 (less positive) and N300 (less negative) components in migrants after~~ a~~ short reimmersion in the L1 context compared to migrants during L2 immersion. Finally, if the decreased access to L1 words primarily affects high-frequency words (as we argue in Casado, Walther, et al., 2023), we should observe larger differences in frequencies for P2 amplitudes in the migrant group after reimmersion in L1 than during immersion in L2. Moreover, we wanted to use the potential differentiation in the effects of P2 and N300 components to shed some light ~~into~~on the time~~-~~ course with which immersion and reimmersion modulate the word production process. If difficulties in picture naming arise during the lexical access stage, we should observe modulations in the naming P2 component. In contrast, if difficulties in picture naming arise during the integration of the lexical information with the perceptual features of the picture, we should observe a modulation in the N300 component.

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

Regarding the first prediction, the results of the present analyses showed no ERP modulation related to overall differences between migrants and control bilinguals. The absence of the overall group effect in the ERPs mirrors the ~~absence~~lack of overall differences between the groups in L1 naming latencies (see Footnote 5 and Casado, Walther, et al., 2023). We, therefore, conclude that despite the long-term immersion of the Polish-English migrants living in the L2 environment (in the UK), we did not find evidence for ~~loss of~~reduced access to the native language, either in behavioural responses~~,~~ or in the electrophysiological markers of the lexical access. Still, the P2 component was sensitive to short-term changes in ~~migrants’~~migrants' language environments, confirming our second prediction. That is, ~~higher~~more positive P2 amplitudes were evoked during L2 immersion vs. after L1 reimmersion, indicating that~~ a~~ short-term reimmersion in the L1 environment resulted in neurocognitive adaptations as reflected in the modulation of the early frontal positivity. Thus, we conclude that the neurocognitive adaptations occur quite early in the process of retrieving the picture's name. It is unclear, however, whether the observed positivity can be identified as the previously reported P2 component~~,~~ (Costa et al., 2009; Strijkers et al., 2010; 2013), as its topographic distribution seems to be more frontal. Finally, we found no evidence for the third prediction, as there was no effect of the lexical frequency; that is, the P2 was insensitive towards any frequency manipulation. The N300 component was also not modulated by any of the variables (migrants vs. control; ~~migrants’~~migrants' L2 immersion vs. L1 reimmersion), thus showing that the effects of immersion and reimmersion during word production ~~can occur quite early in the retrieval of the picture’s~~do not affect the integration process of the visual features of the picture and the lexical information corresponding to the name.

Below, we discuss the implications of our findings for our understanding of the long- and short-term adaptations of bilinguals' cognitive system to the language environment.

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

4.1. Effects of long-term immersion

In line with our behavioural results reported in Casado, Walther et al. (2023), we observed no significant differences in either P2 or N300 mean amplitudes when comparing the evoked responses of migrants with those of the control group. This pattern of results also aligns with previous behavioural results which found that the lexical access in the native language of migrants living in L2 environments for a long time does not differ from bilinguals living in an L1 environment (Beatty-Martinez et al., 2020; Yilmaz & Schmid, 2012) or an environment in which two languages are frequently mixed (Beatty-Martinez et al., 2020). Together with the data reported here, the evidence ~~so far~~ suggests that long-term immersion in the L2 does not necessarily result in generalised difficulty in accessing native language words. Even after ~~a~~ long-term L2 immersion, migrants ~~are able to~~can maintain similar L1 ~~equally accessible~~levels as ~~in~~ bilinguals who live in ~~the~~an L1 environment.

The fact that access to L1 lexical information is not necessarily negatively affected by immersion in the L2 environment (as shown by the absence of group differences in naming latencies and in mean amplitudes) is consistent with the interface hypothesis originally proposed by Sorace and Filiaci (2006). ~~Despite the fact that~~Although their framework ~~offers an explanation for~~explains the difficulties experienced by highly proficient balanced bilinguals, it can also relate to the unbalanced bilinguals with intermediate proficiency, like the participants in the current study. Sorace and Filiaci ~~posit~~posited that long-term L2 immersion only affects the sensitivity to high-level structures external to grammar, such as the interface between syntax and pragmatics that determine appropriateness in context (e.g., during anaphora resolution, see ~~Köpke & Genevska-Hanke, 2018;~~ Chamorro et al., 2016; Chamorro & Sorace, 2016; Köpke & Genevska-Hanke, 2018), rather than causing a permanent change in speakers L1 knowledge representations like

Formatted: Header
Formatted: Font: 12 pt, Italic, Font color: Text 1
Formatted: Justified, Space Before: 14.95 pt, After: 14.95 pt
Formatted: Font: 12 pt, Font color: Text 1
Formatted: Font color: Text 1
Formatted: Justified, Space Before: 12 pt, After: 12 pt
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Footer

Formatted: Header

vocabulary (Chamorro & Sorace, 2019; Sorace, 2011; 2016). Here we provide the first piece of evidence ~~from~~of electrophysiological investigation suggesting that low-level linguistic structures related to the formal semantic features internal to grammatical representations (i.e., lexical level) indeed do not seem to be affected by L2 immersion. ~~This adds~~Thus, considering that the lexical information is part of the lexicon and that we did not find differences in the lexical access abilities between immersed and non-immersed bilinguals, we would propose that long-term L2 immersion does not necessarily modulate the L1 lexicon permanently. In short, these results add to the accumulated body of research indicating that difficulties accessing L1 encountered by bilinguals long-term immersed in the L2 environment affect online sensitivity rather than causing a permanent change in the speaker's L1 knowledge representations (Chamorro et al. 2016; Sorace, 2011; 2016).

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

4.2. Short-term changes triggered by the language environment: L2 immersion vs. L1 reimmersion

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Justified, Space Before: 14.95 pt, After: 14.95 pt

Formatted: Font: 12 pt, Font color: Text 1

Concerning the short-term manipulation of the ~~migrants’~~migrants' language environments, we found that the short-term reimmersion in the L1 environment was related to a change in neural response to picture naming in L1. More specifically, after short-term L1 reimmersion there was a reduction of the positivity at the 160 - 240 ms time window in frontal regions. We intended to explore modulations of the naming P2 component, however previous literature characterised the component slightly differently ~~as~~from the one we found. In particular, the P2 component was previously associated with the ease of lexical access with ~~higher~~more positive amplitudes under more difficult conditions (Costa et al., 2009; Strijkers et al., 2010, 2011~~, 2013~~). Following this

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

interpretation, the decreased amplitude of P2 after L1 reimmersion could indicate that migrants benefited from spending time in the L1 environment, such as this short visit to the native language country facilitated access to L1 compared to during immersion in the L2 environment. However, the interpretation of the P2 component as an index of ease of lexical access was based on its sensitivity to word-lexical frequency, such as low-frequency words evoked ~~higher~~more positive P2 amplitudes ~~compared to~~than high-frequency words (Strijkers et al., 2010). Yet, in the present study, we did not find a modulation of~~-~~ the positivity at the 160~~-~~ - 240 ms time window due to the word-lexical frequency, either as a main effect (which could be expected based on Strijkers et al., 2010) or in an interaction with the environment (unlike in the behavioural analyses reported in Casado, Walther, et al., 2023). Therefore, it is unclear whether the positivity observed in our study is~~ indeed~~ the naming P2 identified in previous research and if it could be interpreted as an index of lexical access. As indicated above, the overall distribution of our ~~fronto-central~~frontocentral component differs from the typical broad ~~centro-parietal~~centroparietal distribution reported in previous studies characterising the P2 as a lexical access index (e.g., Costa et al., 2009; Strijkers et al., 2010; 2011) (see Figure 1 panel ~~A, bottom). Indeed, even~~C). Instead, the ~~study from which we selected this time window and electrode~~ distribution ~~(~~coincides with the more frontal component reported in Strijkers al. (2013). Strijkers et al~~., 2013) shows~~. found a different distribution from the one observed in the current study for the language effect (central, in L1 vs. L2 comparison) and ~~for~~ the lexical frequency (~~centro-parietal~~centroparietal, for low vs. high~~–~~-frequency word comparison). Therefore, it is likely that the positivity in the 160~~-220ms~~ - 220 ms time~~-~~ window found in the present study does not reflect the difficulty of lexical access, but probably a different process, to be better defined.

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

One previous study using a picture-naming task reported a P2 component similar in ~~terms of fronto-central~~ distribution —frontocentral— to the one reported here (Branzi et al., 2014) but offering a different interpretation of the P2 than Strijkers et ~~al's.~~al.'s (2010). Rather than seeing the P2 component as an index of lexical access difficulty, Branzi et al. (2014) suggested that ~~the~~this P2 component could reflect the mechanisms of language ~~regulation~~control; in particular, it could index the application of proactive control. In their study, Branzi et al. (2014) investigated the scope of global-local language control in a group of balanced bilinguals. The participants named new and repeated pictures in their native or second language in blocks divided by language following different orders: L1-L2-L1 or L2-L1-L2. Under the L1-L2-L1 order, the authors found that when balanced bilinguals named pictures in L1 after using L2, ~~higher~~more positive P2 amplitudes were evoked compared to the first block of L1 naming, for both, repeated and new items. The authors argued that the P2 component in their study ~~was reflecting~~reflected language control mechanisms applied during L1 lexical access to ~~proactively~~ manage the persisting activation of the previously used language proactively —in their design, L2— that would create interference within the subsequent naming in the L1.

Following the interpretation of the P2 component proposed by Branzi et al. (2014), our results may indicate that ~~a~~ short-term reimmersion in L1 influences the proactive control mechanisms applied to control for language interference during L1 lexical access. In our study, we found a modulation of the positivity in the 160~~-220ms~~ - 220 ms time-window of the migrant group under the different language environments: the positive amplitudes were decreased after a reimmersion in the L1 environment compared to during immersion in the L2 environment. In our case, the higher positive amplitudes during immersion in the L2 environment could be reflecting the application of proactive control to regulate the interference from the L2 when accessing L1. Similarly, the less

Formatted: Header

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

positive P2 amplitudes after reimmersion in the L1 environment could reflect a decrease in the application of proactive control during L1 access, given the reduced interference from L2 in the L1 environment.

Two other studies support the interpretation suggesting that proactive control plays a crucial role in regulating interference between L1 and L2 in different language environments. Beatty-Martínez et al. (2020) and Zhang et al. (2021) both demonstrated that bilinguals living in their L2 environment, compared to bilinguals in the L1 environment, are more likely to use proactive strategies for cognitive control, measured with the AX-CPT task (Braver et al., 2007). Both studies argue that constant exposure to L2 in the L2 environment trains bilinguals to use alternative control strategies~~ of control~~ to fight interference from the unwanted language (L2 when accessing L1), like proactive control. This interpretation can also be reconciled with the discrepancies between two different P2 distributions observed in comparisons between languages (Strijker et al., 2013) and in comparisons between low- and high-frequency words (Baus et al., 2020; Strijkers et al., 2010, 2011). While both these effects have been previously proposed to reflect lexical access difficulty, they may ~~in fact~~ correspond to two different mechanisms, both relevant to bilingual speech production. The P2 component found in ~~centro-parietal~~centroparietal electrodes (Baus et al., 2020; Strijkers et al., 2010~~, Baus et al., 2020~~) can~~ indeed~~ reflect lexical access difficulty that drives the word-frequency effect. On the ~~other hand~~contrary, the P2 component found over the ~~fronto-central~~frontocentral electrodes may reflect the engagement of proactive control. Engagement of proactive control due to increased interference between languages can be well justified in all studies that report the ~~fronto-central~~frontocentral P2. First, in Strijkers et al. (2013), a ~~larger~~more positive P2 amplitude is observed in speech production in L2, a weaker language that needs to deal with interference from the more strongly activated L1. Second, Branzi et al., (2014)

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

report a ~~larger fronto-central~~more positive frontocentral P2 in response to naming in L1 after L2 compared to a baseline L1 naming. In this case, the use of L2 ~~prior to~~before L1 can lead to a temporary increase in the L2 activation level that results in increased interference between the two languages in the subsequent production ~~in~~of L1 (for the discussion of the influence of recent language use on bilingual language control see Casado, Walther, et al., 2023; Bialystok, 2024; Blanco-Elorrieta & Caramazza, 2021). Finally, our results show ~~increased fronto-central~~more positive frontocentral positivity (P2) in response to speech production during long-term immersion in L2 compared to a context of short-term reimmersion in L1. As previously explained, living in the L2 environment may ~~result in an~~ increase ~~of~~ L2 activation and ~~consequently~~in consequence, it may lead to higher interference between languages.

4.3. The time– course of the role of immersion and reimmersion during word production

We decided to explore the N300 component as it was previously identified as a marker of L1 lexical retrieval difficulties during picture recognition (Wodniecka et al., 2020) and could give us information about whether the possible difficulties in L1 access could be attributed to the integration of visual and lexical information during picture naming at the different environments. We did not find modulations of the mean amplitude of the N300 component, neither in the between-group comparison nor when comparing migrants during L2 immersion and after L1 reimmersion. The lack of N300 modulation in the between-group comparison aligns with the behavioural results (Casado, Walther, et al., 2023), confirming that L1 retrieval abilities during picture recognition were not affected in migrants ~~as a consequence~~because of long-term immersion in the L2 environment. Moreover, the lack of modulation by the context in the migrant group (L2 immersion vs. L1 reimmersion) supports the idea that the positivity between 160 and 240 ms (P2

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Justified, Space Before: 14.95 pt, After: 14.95 pt

Formatted: Font: 12 pt, Italic, Font color: Text 1

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

component) found in this study was not ~~really~~ capturing the ease of lexical access per se, but could have been related to the application of proactive language control during L2 immersion vs. after L1 reimmersion. In sum, the lack of N300 modulations indicates that the changes in the environment did not induce difficulties in the integration during picture recognition.

Altogether, these results could indicate that bilinguals adapt to different language environments by adjusting the language control mechanisms, probably~~-~~ during the lexical processing steps of word production ~~(Indefrey & Levelt, 2004). In order to further test this hypothesis, we~~(Indefrey & Levelt, 2004). We performed an~~-~~ exploratory analysis in which we tested the interaction between Group and Context over a wide range of time windows and distributions to test this hypothesis (see Appendix 6.1). We observed a more widespread scalp distribution and a longer ~~time period~~duration of the ERP modulation due to L1 reimmersion~~, indicating~~. This may indicate that changes in the language environment for the migrant group can be traced through many processes involved during word production and not only ~~-~~related to lexical access. Still, additional studies are required to ~~further~~ explore this finding and its functional significance more deeply.

4.4. ~~Summary~~Other relevant findings

In the P2 analysis, we observed a main effect of age, such that the older the participants, the harder it was to name the pictures as reflected by more positive P2 amplitudes. This result may be explained by two possible sources. On the one hand, older participants experience a decline in domain-general mechanisms including slower processing speed, lower working memory capacity, and decreased inhibitory control (Cabeza et al., 2018; Salthouse, 2010). These aspects also affect specific language processes including word retrieval, which has been shown to decrease RT and increase errors (Burke & Shafto, 2008; Peelle, 2019). On the other hand, previous research showed

Formatted: Header
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font color: Text 1
Formatted: Font: Not Italic, Font color: Text 1
Formatted: Footer

that older participants have greater vocabulary knowledge (Park et al., 2002; Verhaeghen, 2003), which is related to increased interference due to enriched semantic information in older people (Buchler & Reder, 2007; Ramscar et al., 2017). Given that we did not include cognitive tests in our testing battery, we cannot rule out any of the possibilities and further research is needed to clarify it.

Apart from chronological age, we also included L2 proficiency in the analyses as a control variable given that we could not match the participants' group by the General English Test despite our efforts. Our analyses showed that the L2 proficiency did not modulate the ERP response in L1 picture naming. Despite in word recognition studies the modulation of L2 proficiency is expected (e.g., Marian et al, 2003; Spivey & Marian, 1999), it does not seem to be such a clear case for production studies. For instance, Klaus et al. (2018) designed a picture-word interference paradigm in L1 with distractors L2 words presented auditorily, which could be either the translation equivalent of the L1 picture name or unrelated words. They found increased naming latencies when the distractor was the translation equivalent of the to-be-named picture compared to an unrelated distractor, demonstrating that even when speaking in one's dominant L1, translations from the less dominant L2 were active. Importantly, the interference effect was not modulated by L2 proficiency levels. Thus, the naming latencies in L1 did not depend on L2 proficiency. They argue that the L2 proficiency of their participants was rather similar, as in our participants' sample. Further studies should be conducted including a wider L2 proficiency range to explore whether differences in proficiency could modulate L1 naming latencies

Formatted: Header

Formatted: Footer

5. Conclusion

~~In summary, our results reveal that the way in which long-term migrants access L1 words may be different depending on the language environment in which they are currently immersed. Following Branzi et al. (2014)'s~~

In this study, we showed that migrants adapt their neurocognitive system depending on the language environment they are immersed. We observed modulations in the electrophysiological response to naming pictures in the native language, such as more positive amplitudes during L2 immersion and more negative amplitudes after L1 reimmersion in frontal electrodes during the 160 - 240 ms time window (P2 component). Following Branzi et al. (2014)'s interpretation of the P2 component as a marker of proactive control applied by bilinguals to regulate language interference~~,~~ and aligning with Beatty-Martínez et al. (2020) and Zhang et al. (2021) behavioural studies, we propose that the modulation of the early positivity~~ as a function of the short-term changes in the language contexts~~ observed in our study reflects the engagement of proactive control in speech production. Under this explanation, the amount of~~-~~ proactive control applied during immersion in the L2 environment would be higher~~,~~ as the level of L2 activation in this context may be increased. ~~In consequence~~Consequently, controlling the interference between languages in the L2 environment may be more effortful than in the L1 environment. Moreover, the amount of proactive control applied to control for between-language interference would ~~get~~be reduced following ~~a~~ short-term L1 reimmersion. Altogether, our results could be taken as evidence for the hypothesis that the bilingual cognitive system adapts to language environments, adjusting strategies of language control as more or less proactive, depending on the current needs and providing compensatory mechanisms to mitigate the hindering effects of L2 environment on L1 lexical access.

6. ~~5.~~ Limitations and future directions

Formatted: Header

Formatted: Font: Bold, Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

Formatted: Header

The current study is limited by its methodological constraints. Despite trying our best to match the control group with the chronological age, years of education, and L2 proficiency of the migrant group, some differences proved unavoidable, and we could only account for them statistically in the analyses. Moreover, the two groups were tested in different laboratories (one at the University of Edinburgh, and one at the Jagiellonian University in Krakow), which may have induced environmental and technical differences despite the close similarity between labs and equipment. Furthermore, despite our attempts to control for the time passed between their return from holidays in Poland and their testing session "L1 reimmersion", there was some variability between participants, such that some participants were tested the same day and others up to 7 days after. In the latter case, the participants had already been re-exposed to the L2 environment, which might have attenuated some of the effects of L2 reimmersion.

Future studies should replicate the findings with other populations; in particular, it is to be noted that the migrants who participated in our study maintained close contact with the Polish community and culture, as indicated by the relatively high intensity of L1 use. The contact was also maintained by regular trips to their homeland which allowed us to implement the current design. Possibly, this close contact may have induced a protection against reducing the ability to access the native language, which may vary in different populations.

Authors' ~~Conclusion~~

Formatted: Font color: Text 1

~~In this study, we showed that migrants adapt their neurocognitive system depending on the language environment they are immersed in. These changes are manifested by modulations in electrophysiological response to naming pictures in the native language, such as higher amplitudes during L2 immersion and lower amplitudes after L1 reimmersion in frontal electrodes~~

Formatted: Footer

Formatted: Header

~~during the 160-240ms time-window (P2 component). We propose that this pattern of results demonstrates the use of different forms of language control depending on the language environment (more proactive during L2 immersion). Altogether, it seems that the bilingual language system accommodates the long-term experience of using each language and readily adapts to new ratios of each language exposure that changes in each language environment.~~

~~Authors'~~ note

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Justified, Space Before: 16.1 pt, After: 14.95 pt

The data used in the presented models is available here: https://osf.io/bfkp5/?view_only=faab088a3ecd4d9981d3e9e6e636893d.

Formatted: Font color: Text 1

Formatted: Justified

Formatted: Font color: Text 1

The authors declare that they have no conflict of interest.

Formatted: Justified, Space Before: 12 pt, After: 12 pt

Acknowledgments

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Double

The research was possible thanks to a National Science Centre grant (2015/18/E/HS6/00428) awarded to Z.W. and PID2022-142345N-I00 from the Spanish Ministry of Science and Innovation from which A.C. is funded. Special thanks to Dorota Peszkowska and Katarzyna Przybycień, who helped with participant recruitment and testing in Edinburgh. The authors gratefully acknowledge the help of all members of the Psychology of Language and Bilingualism Laboratory, LangUsta, who contributed to the research project by collecting and coding the data, and Michael Timberlake for proofreading.

Formatted: Font color: Text 1

~~6~~7. References

Formatted: Font: 12 pt, Bold, Font color: Text 1

Formatted: Font color: Text 1

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. Health Psychology, 19(6), 586.

Formatted: Justified, Indent: Left: 0", First line: 0", Space Before: 12 pt, After: 12 pt

Formatted: Footer

Ammerlaan, T. (1996). 'You get a bit wobbly...': Exploring bilingual lexical retrieval processes in the context of first language attrition. [Sl: sn].

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models; 2015. Preprint.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. Acta Psychologica, 142(3), 402–409. https://doi.org/10.1016/j.actpsy.2013.01.010

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. Journal of Experimental Psychology. Learning, Memory, and Cognition, 46(6), 1022–1047. https://doi.org/10.1037/xlm0000770

Beatty-Martínez, A. L., & Titone, D. A. (2021). The quest for signals in noise: Leveraging experiential variation to identify bilingual phenotypes. Languages, 6(4), 168.

Bialystok, E. (2024). Bilingualism modifies cognition through adaptation, not transfer. Trends in Cognitive Sciences.

Bice, K., & Kroll, J. F. (2015). Native language change during early stages of second language learning. NeuroReport, 26(16), 966.

Blanco-Elorrieta, E., & Caramazza, A. (2021). A common selection mechanism at each linguistic level in bilingual and monolingual language production. Cognition, 104625. https://doi.org/10.1016/j.cognition.2021.104625

Bogulski, C. A., Bice, K., & Kroll, J. F. (2019). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. Bilingualism: Language and Cognition, 22(5), 1052-1067.

Botezatu, M. R., Kroll, J. F., Trachsel, M. I., & Guo, T. (2021). Second language immersion impacts native language lexical production and comprehension. Linguistic Approaches to Bilingualism.

Branzi, F. M., Martin, C. D., Abutalebi, J., & Costa, A. (2014). The after-effects of bilingual language production. Neuropsychologia, 52, 102-116.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. Variation in working memory, 75, 106.

Formatted: Header

Formatted: Font color: Text 1, English (United Kingdom)

Formatted: Justified, Indent: Left: 0", First line: 0", Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

Buchler, N. E., & Reder, L. M. (2007). Modeling age-related memory deficits: a two-parameter solution. Psychology and aging, 22(1), 104.

Burke, D. M., & Shafto, M. A. (2004). Aging and language production. Current directions in psychological science, 13(1), 21-24.

Cabeza, R., Albert, M., Belleville, S., Craik, F. I., Duarte, A., Grady, C. L., ... & Rajah, M. N. (2018). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. Nature Reviews Neuroscience, 19(11), 701-710.

Casado, A., Walther, J., Wolna, A.,-Szewczyk, J. , Sorace, A., & Wodniecka, Z. (2022). Advantages of coming back: how reimmersion in the native country impacts ~~migrants’~~migrants' native language access. [manuscript in review] Institute of Psychology, Jagiellonian University.

Chamorro, G., Sorace, A., & Sturt, P. (2016). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition*. Bilingualism: Language and Cognition, 19(3), 520–532. https://doi.org/10.1017/S1366728915000152

Chamorro, G., & Sorace, A. (2019). The Interface Hypothesis as a framework for studying L1 attrition. In The Oxford handbook of language attrition (pp. 24-35). Oxford University Press.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical Selection in Bilinguals: Do Words in the Bilingual's Two Lexicons Compete for Selection? Journal of Memory and Language, 41(3), 365–397. https://doi.org/10.1006/jmla.1999.2651

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. Proceedings of the National Academy of Sciences, 106(50), 21442-21446.

Delorme A & Makeig S (2004) EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics, Journal of Neuroscience Methods 134:9-21

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage, 34(4), 1443-1449.

Ecke, P., & Hall, C. J. (2013). Tracking tip-of-the-tongue states in a multilingual speaker: Evidence of attrition or instability in lexical systems? International Journal of Bilingualism, 17(6), 734–751. https://doi.org/10.1177/1367006912454623

Finkbeiner, M., Almeida, J., Janssen, N., & Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(5), 1075–1089. https://doi.org/10.1037/0278-7393.32.5.1075

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. Behavior Research Methods, Instruments, & Computers, 35(1), 116–124. https://doi.org/10.3758/BF03195503

Formatted: Header

Formatted: Font color: Text 1, English (United Kingdom)

Formatted: Justified, Indent: Left: 0", First line: 0", Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

Gargiulo, C., & van de Weijer, J. (2020). Anaphora resolution in L1 Italian in a Swedish-speaking environment before and after L1 re-immersion: A study on attrition. Lingua, 233, 102746. https://doi.org/10.1016/j.lingua.2019.102746

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. Journal of Memory and Language, 58(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. Bilingualism: Language and Cognition, 1(2), 67–81. https://doi.org/10.1017/S1366728998000133

Haman, E., Łuniewska, M., Hansen, P., Simonsen, H. G., Chiat, S., Bjekić, J., … Armon-Lotem, S. (2017). Noun and verb knowledge in monolingual preschool children across 17 languages: Data from Cross-linguistic Lexical Tasks (LITMUS-CLT). Clinical Linguistics & Phonetics, 31(11–12), 818–843. https://doi.org/10.1080/02699206.2017.1308553

Haman, E., Łuniewska, M., & Pomiechowska, B. (2015). 8. Designing Cross-Linguistic Lexical Tasks (CLTs) for Bilingual Preschool Children. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), Assessing Multilingual Children (pp. 196–240). Multilingual Matters. https://doi.org/doi:10.21832/9781783093137-010

Hulsen, M. E. H. (2000). Language loss and language processing: Three generations of Dutch migrants in New Zealand. [S.l. : s.n.]. ~~Retrieved from https://repository.ubn.ru.nl/handle/2066/18901~~Retrieved from https://repository.ubn.ru.nl/handle/2066/18901

Indefrey, P., and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144. doi: 10.1016/j.cognition.2002.06.001

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clinical Neurophysiology, 111(10), 1745-1758.

Klaus, J., Lemhöfer, K., & Schriefers, H. (2018). The second language interferes with picture naming in the first language: Evidence for L2 activation during L1 production. Language, Cognition and Neuroscience, 33(7), 867-877.

Köpke, B., & Genevska-Hanke, D. (2018). First Language Attrition and Dominance: Same or Different? Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.01963

Köpke, B., & Schmid, M. S. (2004). Language attrition. First Language Attrition: Interdisciplinary Perspectives on Methodological Issues, 28(1).

Kuznetsova A, Brockhoff PB, Christensen RHB (2017). ~~“~~"lmerTest Package: Tests in Linear Mixed Effects Models~~.”~~." *Journal of Statistical Software*, **82**(13), 1–26~~.~~ ~~doi:10.18637/jss.v082.i13.~~.doi:10.18637/jss.v082.i13.

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Indent: Left: 0", First line: 0", Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Footer

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. Behavior research methods, 44(2), 325-343.

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. Journal of Statistical Software, 69(1), 1–33. ~~https://doi.org/10.18637/jss.v069.i01~~https://doi.org/10.18637/jss.v069.i01

Lenth, R. V. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. ~~https://CRAN.R-project.org/package=emmeans~~https://CRAN.R-project.org/package=emmeans

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing Access to the Native Language While Immersed in a Second Language: Evidence for the Role of Inhibition in Second-Language Learning. Psychological Science, 20(12), 1507–1515. https://doi.org/10.1111/j.1467-9280.2009.02480.x

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. Frontiers in human neuroscience, 8, 213.

Mandera, P., Keuleers, E., Wodniecka, Z., & Brysbaert, M. (2015). Subtlex-pl: Subtitle-based word frequency estimates for Polish. Behavior Research Methods, 47(2), 471–483. https://doi.org/10.3758/s13428-014-0489-4

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. Brain and language, 86(1), 70-82.

Park, D. C. (2002). Aging, cognition, and culture: a neuroscientific perspective. Neuroscience & Biobehavioral Reviews, 26(7), 859-867.

Peelle, J. E. (2019). Language and aging. The Oxford handbook of neurolinguistics, 10, 1-21.

Ramscar, M., Sun, C. C., Hendrix, P., & Baayen, H. (2017). The mismeasurement of mind: Life-span changes in paired-associate-learning scores reflect the "cost" of learning, not cognitive decline. Psychological science, 28(8), 1171-1179.

Salthouse, T. A. (2010). Selective review of cognitive aging. Journal of the International neuropsychological Society, 16(5), 754-760.

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. Bilingualism: Language and Cognition, 13(2), 231-252.

Schmid, M. S., & Jarvis, S. (2014). Lexical access and lexical diversity in first language attrition. Bilingualism: Language and Cognition, 17(4), 729–748. https://doi.org/10.1017/S1366728913000771

Schmid, M. S., & Keijzer, M. (2009). First language attrition and reversion among older migrants. International Journal of the Sociology of Language, 2009(200), 83–101. https://doi.org/10.1515/IJSL.2009.046

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1, English (United Kingdom)

Formatted: Font color: Text 1

Formatted: Underline, Underline color: Custom Color(RGB(0,0,238)), Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Indent: Left: 0", First line: 0", Space Before: 12 pt, After: 12 pt

Formatted: Footer

Sorace, A. (2011). Pinning down the concept of “"interface”" in bilingualism. Linguistic approaches to bilingualism, 1(1), 1-33.

Sorace, A. (2016). Referring expressions and executive functions in bilingualism. Linguistic Approaches to Bilingualism, 6(5), 669–684.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. Psychological science, 10(3), 281-284.

Strijkers, K., Baus, C., Runnqvist, E., FitzPatrick, I., & Costa, A. (2013). The temporal dynamics of first versus second language production. Brain and language, 127(1), 6-11.

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. Cerebral cortex, 20(4), 912-928.

Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. Journal of memory and language, 65(4), 345-362.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. Proceedings of the National Academy of Sciences, 104(30), 12530–12535. https://doi.org/10.1073/pnas.0609927104

Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. International Journal of Bilingualism, 8(3), 257–277. https://doi.org/10.1177/13670069040080030601

Valente, A., Bürki, A., & Laganaro, M. (2014). ERP correlates of word production predictors in picture naming: a trial by trial multiple regression analysis from stimulus onset to response. Frontiers in neuroscience, 8, 390.

Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. Psychology and aging, 18(2), 332.

Wodniecka, Z., Szewczyk, J., Kałamała, P., Mandera, P., & Durlik, J. (2020). When a second language hits a native language. What ERPs (do and do not) tell us about language retrieval difficulty in bilingual language production. Neuropsychologia, 141, 107390.

Wolna, A., Łuniewska, M., Haman, E., & Wodniecka, Z. (2023). Polish norms for a set of colored drawings of 168 objects and 146 actions with predictors of naming performance. Behavior Research Methods, 55(5), 2706-2732.

Yagmur, K., Bot, K. D., & Korzilius, H. (1999). Language Attrition, Language Shift and Ethnolinguistic Vitality of Turkish in Australia. Journal of Multilingual and Multicultural Development, 20(1), 51–69. https://doi.org/10.1080/01434639908666369

Formatted: Header

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Indent: Left: 0", First line: 0", Space Before: 12 pt, After: 12 pt

Formatted: Underline, Underline color: Custom Color(RGB(0,0,238)), Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1, Polish

Formatted: Justified, Indent: Left: 0", First line: 0", Space Before: 12 pt, After: 12 pt

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Indent: Left: 0", First line: 0", Space Before: 12 pt, After: 12 pt

Formatted: Footer

Yilmaz, G., & Schmid, M. S. (2012). L1 accessibility among Turkish-Dutch bilinguals. The Mental Lexicon, 7(3), 249–274. https://doi.org/10.1075/ml.7.3.01yil

Zhang, H., Diaz, M. T., Guo, T., & Kroll, J. F. (2021). Language immersion and language training: Two paths to enhanced language regulation and cognitive control. Brain and Language, 223, 105043.

Formatted: Header

Formatted: Footer

Formatted: Header

~~7~~

8 . Appendices

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Justified, Space Before: 16.1 pt, After: 16.1 pt, Line spacing: Double

A. Exploratory analysis

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Double

Formatted: Font: Not Italic, Font color: Text 1

~~In order to~~To check if the effect of L1 reimmersion was limited to the positivity, investigated in the a priori analysis, we conducted an exploratory analysis. We divided the data into five time windows; 0–60 ms, 60-140 ms, 140–250 ms, 250-350 ms, and 350–500 ms. Additionally, we used ten electrode cluster: Frontal (Fp1, AF3, AF4, Fp2), LeftAnterior (F7, F3, FC5), Fronto-central (Fz, FC1, FC2), RightAnterior (F8, F4, FC6), Left-central (C3, CP5), Centro-Parietal (Cz, CP1, CP2, Pz), Right-central (C4, CP6), Left-parietal (P7, P3, PO3), Right-parietal (P8, P4, PO4), Occipital (O1, Oz, O2).

Formatted: Font color: Text 1

Formatted: Font color: Text 1

A 2 × 2 within-subjects ANOVA was conducted considering Group and Context in all time windows and scalp areas.

Formatted: Justified, Space Before: 12 pt, Line spacing: Double

Table ~~A.1~~A.1 shows the significance of the Group:Context interaction for each time~~-~~ window and electrode cluster. The effect seems to arise already during the earliest time window and is more pronounced on electrodes on the left hemisphere. Overall though the Group:Context interaction still ~~seems to~~ have a broad spread over most scalp areas and time windows. Therefore, it seems that the change in the language environment can be traced through many processes involved during word production, from early processing steps to later ones around 350ms.

Formatted: Justified, Line spacing: Double

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Table A.1

Formatted: Justified, Space Before: 12 pt, Line spacing: Double

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Footer

Statistics resulting from the ANOVAs conducted at each temporal window and scalp area for the Group:Context interaction.

Electrode cluster	0-60	60-140	140-250	250-350	350-500
Frontal	$F_{(1,1)} = 3.53$, $p = 0.07$	$F_{(1,1)} = 0.94$, $p = 0.34$	$F_{(1,1)} = 1.17$, $p = 0.28$	$F_{(1,1)} = 2.33$, $p = 0.13$	$F_{(1,1)} = 2.10$, $p = 0.15$
LeftAnterior	**$F_{(1,1)} = 6.20$, $p = 0.02$**	**$F_{(1,1)} = 5.32$, $p = 0.03$**	**$F_{(1,1)} = 5.99$, $p = 0.02$**	**$F_{(1,1)} = 7.03$, $p = 0.01$**	**$F_{(1,1)} = 4.90$, $p = 0.03$**
Fronto-central	$F_{(1,1)} = 3.73$, $p = 0.06$	**$F_{(1,1)} = 5.49$, $p = 0.02$**	**$F_{(1,1)} = 6.43$, $p = 0.01$**	**$F_{(1,1)} = 4.50$, $p = 0.04$**	$F_{(1,1)} = 1.36$, $p = 0.25$
RightAnterior	$F_{(1,1)} = 1.60$, $p = 0.22$	$F_{(1,1)} = 0.39$, $p = 0.53$	$F_{(1,1)} = 2.08$, $p = 0.16$	$F_{(1,1)} = 1.09$, $p = 0.30$	$F_{(1,1)} = 1.49$, $p = 0.23$
Left-central	**$F_{(1,1)} = 6.83$, $p = 0.01$**	**$F_{(1,1)} = 5.48$, $p = 0.02$**	**$F_{(1,1)} = 5.99$, $p = 0.02$**	$F_{(1,1)} = 3.91$, $p = 0.05$	$F_{(1,1)} = 2.23$, $p = 0.14$
Centro-Parietal	**$F_{(1,1)} = 8.17$, $p = 0.01$**	**$F_{(1,1)} = 5.52$, $p = 0.02$**	**$F_{(1,1)} = 4.75$, $p = 0.03$**	$F_{(1,1)} = 1.86$, $p = 0.18$	$F_{(1,1)} = 1.33$, $p = 0.25$
Right-central	**$F_{(1,1)} = 5.24$, $p = 0.03$**	$F_{(1,1)} = 2.46$, $p = 0.12$	**$F_{(1,1)} = 6.28$, $p = 0.02$**	$F_{(1,1)} = 1.31$, $p = 0.26$	$F_{(1,1)} = 1.03$, $p = 0.31$
Left-parietal	**$F_{(1,1)} = 13.98$, $p < 0.001$**	**$F_{(1,1)} = 6.18$, $p = 0.02$**	**$F_{(1,1)} = 4.12$, $p = 0.05$**	$F_{(1,1)} = 2.68$, $p = 0.11$	$F_{(1,1)} = 2.41$, $p = 0.13$
Right-parietal	**$F_{(1,1)} = 8.90$, $p < 0.01$**	$F_{(1,1)} = 1.35$, $p = 0.25$	$F_{(1,1)} = 1.46$, $p = 0.23$	$F_{(1,1)} = 0.88$, $p = 0.35$	$F_{(1,1)} = 1.48$, $p = 0.23$
Occipital	**$F_{(1,1)} = 10.86$, $p < 0.01$**	$F_{(1,1)} = 3.92$, $p = 0.05$	$F_{(1,1)} = 3.13$, $p = 0.08$	$F_{(1,1)} = 2.77$, $p = 0.10$	$F_{(1,1)} = 1.32$, $p = 0.26$

Formatted: Header

Formatted: Font color: Text 1

Formatted: Justified, Line spacing: Double

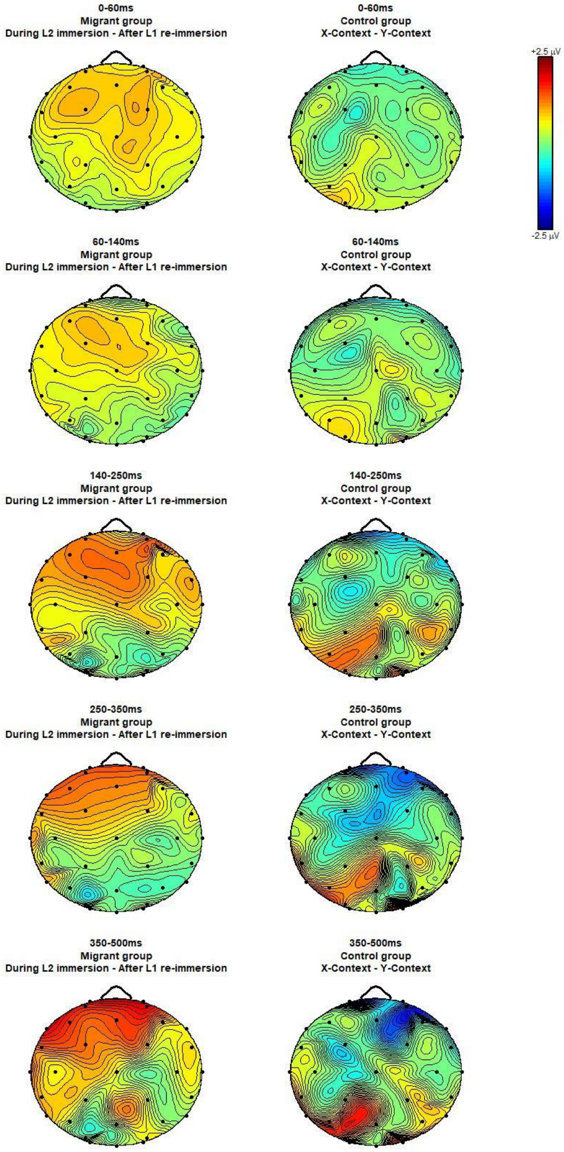
Formatted: Font: Not Bold, Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Double

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Footer

Figure A.1



Formatted: Header

Formatted: Font color: Text 1

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Footer

Note. Scalp topography maps for all time windows of the exploratory analysis for the migrant group (left column) and the control group (right column).

B. Gender analysis

To ensure that our analysis was not affected by the disproportionate gender distribution in our participant sample, we ~~conducted an analysis of~~analysed the ERP data involving gender as a fixed factor. Its inclusion did not significantly change our results with the P2 model (Table B.1) retaining its significant interaction between Group and Context. In the same line, there were no changes for the N300 model (Table B.2). ~~Together~~Altogether, this indicates that gender did not play a significant role in how the language system adapts to changes in the language environment. However, given the uneven gender distribution in our sample, more studies involving a more balanced participant group are still needed.

Table B.1

Fixed effects for the LME model with ERP amplitude and both groups for the P2 component including Gender as a fixed factor.

	Estimate	SE	t	by-Picture SD	by-Participant SD
Intercept	-0.06	0.08	-0.69	0.10	0.31
Group	-0.09	0.11	-0.87		-
Context	0.03	0.03	0.94		0.16
Word-lexical frequency	0.01	0.04	0.24	-	

Formatted: Header

Formatted: Font: 11 pt, Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Multiple 1.15 li

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Double

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, Line spacing: Double

Formatted: Font: Not Italic, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: Not Italic, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Justified, Line spacing: Double

Formatted Table

Formatted: Footer

Participant's Gender	-0.11	0.11	-1.06		-
Participant's Age	0.23	0.05	4.84***		-
Age of L2 acquisition	-0.08	0.05	-1.83		-
L2 proficiency	-0.06	0.04	-1.34		-
log (Trial number)	0.01	0.02	0.40		
Group: Context	0.17	0.06	2.73**		-
Group: Word-lexical frequency	0.13	0.08	1.57	-	-
Control Group : AfterL1-Immersion : Word-lexical frequency : Female	0.05	0.05	0.88	-	-
Migrant Group : AfterL1-Immersion : Word-lexical frequency : Female	-0.09	0.08	-1.24	-	-
Control Group : DuringL2-Immersion : Word-lexical frequency : Female	0.09	0.05	1.78	-	-
Experimental Group : DuringL2-Immersion : Word-lexical frequency : Female	-0.06	0.08	-0.84	-	-
Control Group : AfterL1-Immersion : Word-lexical frequency : Male	0.08	0.06	1.41	-	-
Experimental Group : AfterL1-Immersion : Word-lexical frequency : Male	-0.06	0.10	-0.54	-	-

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.2

Formatted: Header

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Justified, Line spacing: Double

Formatted: Font color: Text 1

Formatted: Justified, Space Before: 12 pt, Line spacing: Double

Formatted: Font: Not Bold, Font color: Text 1

Formatted: Footer

Fixed effects for the LME model with ERP amplitude and both groups for the N300 component including Gender as a fixed factor.

	Estimate	SE	t	by-Picture SD	by- Participant SD
Intercept	-0.08	0.09	-0.89	0.23	0.39
Group	-0.20	0.13	-1.51		-
Context	0.02	0.04	0.69	0.04	0.20
Word-lexical frequency	0.03	0.04	0.70	-	
Participant's Gender	-0.05	0.13	-0.40		-
Participant's Age	0.07	0.06	1.17		-
Age of L2 acquisition	-0.07	0.06	-1.21		-
L2 proficiency	0.01	0.05	0.27		-
log (Trial number)	0.02	0.02	1.15		
Group: Context	0.11	0.07	1.54		-
Group: Word-lexical frequency	0.06	0.08	0.69	-	-
Control Group : AfterL1-Immersion : Word-lexical frequency : Female	0.01	0.05	0.10	-	-
Migrant Group : AfterL1-Immersion : Word-lexical frequency : Female	-0.06	0.07	-0.80	-	-

Formatted: Header

Formatted: Font color: Text 1

Formatted: Justified, Line spacing: Double

Formatted Table

Formatted: Footer

Control Group : DuringL2-Immersion : Word-lexical frequency : Female	0.03	0.05	0.57	-	-
Experimental Group : DuringL2-Immersion : Word-lexical frequency : Female	-0.03	0.07	-0.35	-	-
Control Group : AfterL1-Immersion : Word-lexical frequency : Male	0.04	0.06	0.80	-	-
Experimental Group : AfterL1-Immersion : Word-lexical frequency : Male	0.05	0.10	0.50	-	-

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Formatted: Header

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Justified, Space Before: 12 pt, After: 12 pt, Line spacing: Double

Formatted: Footer

Figure 1

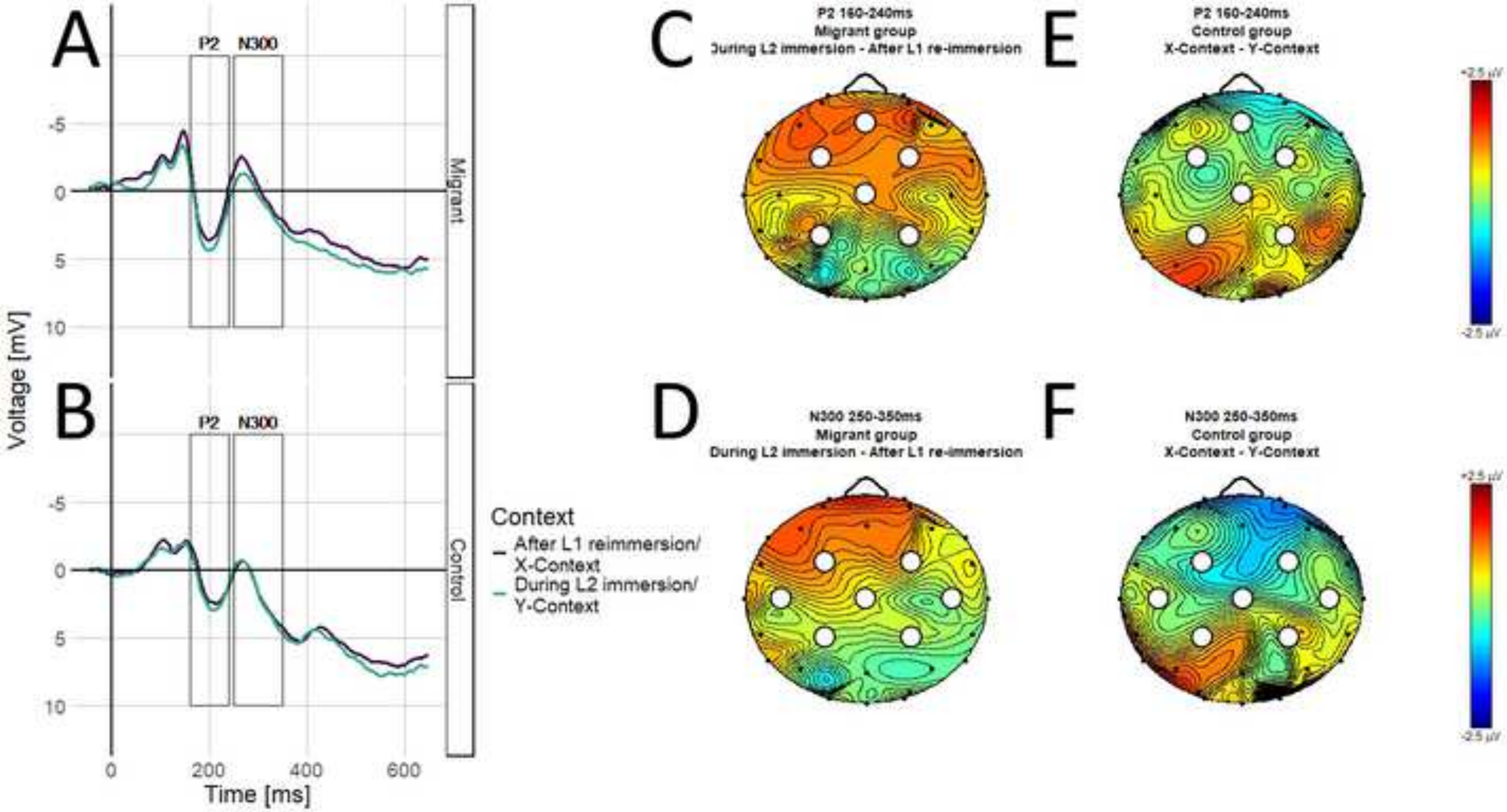

A
P2
N300
-5
0
5
10
Migrant
B
P2
N300
-5
0
5
10
Control
Voltage [mV]
0
200
400
600
Time [ms]
Context
After L1 reimmersion/ X-Context
During L2 immersion/ Y-Context
C
P2 160-240ms
Migrant group
During L2 immersion - After L1 re-immersion
E
P2 160-240ms
Control group
X-Context - Y-Context
+2.5 µV
-2.5 µV
D
N300 250-350ms
Migrant group
During L2 immersion - After L1 re-immersion
F
N300 250-350ms
Control group
X-Context - Y-Context
+2.5 µV
-2.5 µV

Table 1

Demographic information and language experience of participants.

	Migrant group (N = 32)		Control group (N = 32)		t-test	
N	32 (29 female)		32 (23 female)			
Participant’s Age (years)	36.16 [33.83, 38.48]		29.84 [27.13, 32.56]		t(62) = -4.19, p = <0.01***	
SES (1 - 10)	6.64 [6.09, 7.20]		5.91 [5.33, 6.48]		t(62) = -1.62, p = 0.11	
Years of education	18.35 [17.43, 19.28]		17.16 [16.39, 17.92]		t(62) = -2.34, p = 0.02*	
Length of residence in L2 environment (years)	9.66 [7.90, 11.41]		-			
Length of reimmersion in L1 environment (days)	13.38 [10.42, 16.33]		-			
Time delay between L1 reimmersion and recording (days)	3.06 [2.39, 3.74]		-			
Self-assessed language experience (1-10)	L1	L2	L1	L2	L1	L2
Self-rated proficiency	9.82 [9.65, 9.99]	7.86 [7.52, 8.19]	9.71 [9.46, 9.96]	7.17 [6.74, 7.60]	t(62) = -0.57, p = 0.57	t(62) = -2.33, p = 0.02*
Speaking	9.64 [9.36, 9.92]	7.68 [7.23, 8.13]	9.56 [9.17, 9.95]	6.59 [6.10, 7.09]	t(62) = -0.20, p = 0.84	t(62) = -3.02, p = <0.01*
Writing	9.77 [9.50, 10.04]	7.42 [6.96, 7.89]	9.66 [9.40, 9.91]	6.66 [6.04, 7.27]	t(62) = -0.28, p = 0.78	t(62) = -1.74, p = 0.09

Listening	9.94 [9.85, 10.02]	7.91 [7.60, 8.22]	9.84 [9.66, 10.03]	7.53 [7.12, 7.94]	t(62) = -0.91,	t(62) = -1.37,
					p = 0.37	p = 0.17
Reading	9.93 [9.85, 10.02]	8.42 [8.07, 8.77]	9.78 [9.50, 10.07]	7.91 [7.43, 8.39]	t(62) = -1.05,	t(62) = -1.74,
					p = 0.30	p = 0.09
Percentage of daily use (%)	40.46 [34.72, 46.21]	59.25 [53.66, 64.85]	81.32 [75.44, 87.19]	16.84 [12.75, 20.93]	t(62) = 10.29,	t(62) = -12.68,
					p = <0.01***	p = <0.01***
Age of L2 acquisition (years)	-	13.05 [11.70, 14.39]	-	10.00 [8.41, 11.59]	t(62) = -3.89, p = <0.01***	
Intensity of language switching (1 = no switching, 10 = always switching)	4.82 [3.90, 5.74]		4.06 [3.27, 4.85]		t(62) = -0.84, p = 0.40	
Objective L2 proficiency measures						
LexTALE (mean accuracy in %)	-	0.78 [0.73, 0.83]	-	0.74 [0.70, 0.77]	t(62) = -1.55, p = 0.13	
General English Test (mean accuracy in %)	-	89.88 [86.44, 93,28]	-	85.12 [82.44, 87.84]	t(62) = -2.54, p = 0.01*	

Note. The first part of the table describes the demographic information of the final migrant group and the final control group. The rows display (1) the number of participants with the number of women in brackets (gender was self-reported by participants), (2) age (in years), (3) socio-economic status on a 1 to 8 scale based on Adler et al. (2000), (4) years of education (in years), (5) length of residence in an L2 environment (in years), (6) length of immersion the L1 environment and (7) time passed between arrival at the airport in the UK and the experimental recording session. The second part of the table summarises the self-assessed language experience based on a questionnaire. The self-rated proficiency is presented on a scale from 1 to 10, where 1 = "no knowledge of a given language" and 10 = "native-like proficiency". The daily use of each language is presented in percentages while the age of acquisition is presented in years. Bilingual switching is presented on a scale from 1 to 10, where 1= "I never switch languages within sentences" and 10 = "I always switch languages within sentences". The objective L2 proficiency measures in English are presented in percentages. 95% confidence intervals are given in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2

Stimuli information for the four subsets used during picture naming.

	A	B	C	D
PL Age of Acquisition (years)	3.43	3.56	3.41	3.49
	[3.27, 3.59]	[3.40, 3.73]	[3.28, 3.54]	[3.34, 3.64]
Name agreement (%)	0.93	0.92	0.95	0.93
	[0.89, 0.96]	[0.88, 0.95]	[0.92, 0,97]	[0.90, 0.96]
Lexical frequency	2.77	2.71	2.72	2.85
	[2.59, 2.95]	[2.55, 2.88]	[2.54, 2.89]	[2.66, 3.04]
Phoneme length	5.58	5.38	5.72	5.60
	[5.20, 5.96]	[4.94, 5.82]	[5.32, 6.13]	[5.15, 6.05]

Note. The table shows the average stimuli properties of the used picture subsets. The rows display (1) the age of acquisition for the Polish picture names (in years) based on Haman et al. (2015), (2) the name agreement of the used pictures (in percent) based on Wolna et al. (2022), (3) lexical frequency based on Mandera et al. (2015), (4) phoneme length and (5) Imageability based on Wolna et al. (2022). 95% confidence intervals are given in parentheses.

Table 3

Summary of the final trial count after artefact rejection during EEG processing and removal of inaccurate responses and too early articulator movements.

	During L2 immersion/X-Context	After L1 reimmersion/Y-Context
Migrant group	88.6% (1531 trials)	82.0% (1417 trials)
Control group	87.0% (1504 trials)	85.6% (1479 trials)

Note. The table gives the percentage and number (in brackets) of remaining trials in the final analysis for each Group and Context. The raw data set contained 1728 trials within each condition.

Table 4

Fixed effects for the LME model with ERP amplitude and both groups for the P2 component.

	Estimate	SE	t	by- Picture SD	by- Participant SD
Intercept	-0.03	0.08	-0.44	0.10	0.31
Group	-0.05	0.10	-0.47		-
Context	0.03	0.03	0.86		0.16
Word-lexical frequency	0.00	0.01	0.22	-	
Participant’s Age	0.22	0.05	4.71***		-
Age of L2 acquisition	-0.09	0.05	-1.91		-
L2 proficiency	-0.07	0.04	-1.58		-
log (Trial number)	0.01	0.02	0.54		
Group: Context	0.18	0.06	2.76**		-
Group: Word-lexical frequency	0.00	0.02	-0.12	-	-
Control Group: Context: Word-lexical frequency	0.00	0.03	0.08	-	-
Migrant Group: Context: Word-lexical frequency	0.03	0.03	0.89	-	-

Table 5

Fixed effects for the LME model with ERP amplitude and both groups for the N300 component.

	Estimate	SE	t	by- Picture SD	by- Participant SD
Intercept	-0.09	0.09	-0.98	0.23	0.39
Group	-0.18	0.12	-1.50		-
Context	0.02	0.04	0.63		0.20
Word-lexical frequency	0.02	0.02	1.15	-	
Participant’s Age	0.06	0.06	1.10		-
Age of L2 acquisition	-0.07	0.06	-1.26		-
L2 proficiency	0.01	0.05	0.21		-
log (Trial number)	0.03	0.02	1.35		
Group: Context	0.11	0.07	1.52		-
Group: Word-lexical frequency	0.00	0.02	-0.05	-	-
Control Group: Context: Word-lexical frequency	0.00	0.03	-0.10	-	-
Migrant Group: Context: Word-lexical frequency	0.02	0.03	0.61	-	-

Table 6

Summary of the electrophysiological data.

ERP component	Group	During L2 immersion / X-Context	After L1 reimmersion / Y-Context	Context comparison
P2	Migrant group	0.04 (0.07)	-0.08 (0.06)	*z* = -2.59; *p* = 0.01*
	Control group	-0.01 (0.07)	0.05 (0.06)	*z* = 1.33; *p* = 0.19
	Group comparison	*z* = -0.40; *p* = 0.69	*z* = 1.40; *p* = 0.16	
N300	Migrant group	-0.04 (0.08)	-0.11 (0.08)	*z* = -1.55; *p* = 0.12
	Control group	0.08 (0.08)	0.11 (0.08)	*z* = 0.62; *p* = 0.53
	Group comparison	*z* = 1.00; *p* = 0.32	*z* = 1.87; *p* = 0.06	

Note. The table gives the predicted amplitudes of the two ERP time windows for the two groups in the two Contexts and the pairwise comparison from their respective linear-mixed models using the emmeans function. Standard errors are given in parentheses.

18/06/2024

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The author is an Editorial Board Member/Editor-in-Chief/Associate Editor/Guest Editor for *[Journal name]* and was not involved in the editorial review or the decision to publish this article.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Alba Casado on behalf of other authors

Author’s contribution statement

(based on CRediT Taxonomy)

Credit Statement to the publication: Casado, A., Walther, J., Wolna, A., Szewczyk, J., Sorace, A., & Wodniecka, Z. (2024). Investigation of long- and short-term adaptations of the bilingual language system to different language environments: evidence from the ERPs

Name	Conceptualization	Methodology	Software	Validation	Formal analysis	Investigation	Resources	Data curation	Writing (Original Draft)	Writing (Review & Editing)	Visualization	Supervision	Project administration	Funding acquisition
Jonas Walther	X	X	X	X	X			X	X	X	X			
Alba Casado	X	X	X	X	X				X	X	X			
Agata Wolna	X	X	X			X				X				
Jakub Szewczyk	X	X								X				
Antonella Sorace	X						X							
Zofia Wodniecka	X	X					X			X		X	X	X

Figure captions