DOCTORAL THESIS
Institute of Psychology
Jagiellonian University

REGULATION, COMPETITION, OR COOPERATION? EXPLORING LANGUAGE ACCESS AND COGNITIVE CONTROL UNDER DIFFERENT ENVIRONMENTAL DEMANDS.

Jonas Walther

Thesis prepared under the supervision of
dr. hab. Zofia Wodniecka-Chlipalska, prof. UJ
dr. Kalinka Timmer

Krakow, 2023

LANGUSTA
Psychology of Language and Bilingualism Lab

JAGIELLONIAN UNIVERSITY
IN KRAKÓW

Declaration

I hereby declare that this thesis represents my own work which has been done after registration for the degree of PhD at Jagiellonian University and has not been previously included in a thesis or dissertation submitted to this or any other institution for a degree, diploma or other qualifications. The co-authors of the articles indicated and confirmed their contribution to the presented investigations. The authors' contribution confirmations are attached to the doctoral dissertation.

Jonas Walther

July 3, 2023

Funding

This project has received funding from Sonata BIS grant 015/18/E/HS6/00428 from National Science Centre awarded to Zofia Wodniecka.

With special thanks to:

Prof. Zofia Wodniecka and dr. Kalinka Timmer, for their invaluable guidance and encouragement over the years;

Dr. Patrycja Kałamała and Dr. Alba Casado, for their mentorship, patience and support;

All Langusta's members, for all the feedback and the time spent together;

My parents and brother, for their support in this endeavor.

Table of contents

Abstract

The so-called "bilingual advantage" – the potential benefits for domain-general cognitive control as the result of bilingualism, has not been consistently replicated, and researchers have developed different theories to explain this variation in results. One explanation is that bilingualism should not be treated as a monolithic factor; instead, each bilingual's experiences and individual differences should be considered, as well as how language is used in different environments. Three broad language contexts have been suggested: regulatory, competitive, and cooperative. This dissertation explores whether these language environments differentially impact language use and cognitive control. Investigations 1 and 2 studied language availability and language control mechanisms in regulative and competitive language contexts. Investigation 3, on the other hand, attempted to simulate the competitive and cooperative contexts within a short-term experimental study and to observe the effects they may exert on specific domain-general cognitive mechanisms. By using a combination of behavioral and electrophysiological measures and restricting the scope to word production in Polish-English bilinguals, this dissertation attempts to give a broader understanding of the relationship between language environment and its consequences on the cognitive system. Additionally, the use of within-group and between-group comparisons in each study allowed for a more comprehensive exploration of changes in the cognitive system due to changes in the language environment.

Collectively, this dissertation aims to advance our understanding of the cognitive effects of bilingualism and to shed light on the flexibility of the language system. Investigations 1 and 2 found that the native language of long-term migrants is sensitive to changes in the environment. Even a short-term reimmerison in L1 results in a change in language control strategies. Investigation 3 found that the short-term language switching training, designed to simulate either competitive or cooperative language use, also results in a change of cognitive control processes of bilinguals, leading them to rely on more global monitoring in a competitive language context. In summary, the dissertation paints a picture of a fast and adaptive language system that rapidly changes its cognitive strategies based on the environment's demands. The findings from the three presented investigations contribute to the existing body of knowledge in psychology, providing insights into the mechanisms underlying bilingualism-related cognitive benefits.

Keywords: bilingualism, cognitive control, language context, lexical access, cognitive demand

Streszczenie

Tak zwana "dwujęzyczna przewaga" - potencjalne korzyści dla ogólnej kontroli poznawczej w wyniku dwujęzyczności, nie replikuje się we wszystkich badaniach, a naukowcy próbują zrozumieć i wyjaśnić rozbieżności w wynikach badań. zmienność wyników. Jednym z wyjaśnień jest to, że dwujęzyczność nie jest zjawiskiem jednorodnym i wymaga uwzględnienia złożoności i różnorodności doświadczeń osób dwujęzycznych, a także różnorodności środowisk językowych. Przebadano trzy szerokie konteksty językowe: regulacyjny, konkurencyjny i kooperacyjny. W ramach rozprawy przeanalizowano a, w jaki sposób te trzy środowiska językowe wpływają na używanie języka i kontrolę poznawczą. W Badaniach 1 i 2 badano dostęp leksykalny i mechanizmy kontroli językowej w regulacyjnym i konkurencyjnym kontekście językowym. Z kolei w Badaniu 3 podjęto próbę symulacji kontekstu rywalizacyjnego i kooperacyjnego w ramach badania eksperymentalnego, obserwując wpływ, jaki każdy z kontekstówwywiera na mechanizmy poznawcze. Wykorzystując kombinację pomiarów behawioralnych i elektrofizjologicznych oraz koncentrując sięna badaniu produkcji słów u dwujęzycznych osób polsko-angielskich, starałem się lepiej zrozumić konsekwencje środowiska językowego nadziałanie systemu poznawczego. Zastosowanie porównań wewnątrzgrupowych i międzygrupowych w każdym badaniu dało możliwość lepszej kontroli czynników zakłócających.

Podsumowując, niniejsza rozprawa ma na celu pogłębienie naszego zrozumienia poznawczych skutków dwujęzyczności i rzucenie światła na elastyczność systemu językowego. Badania 1 i 2 wykazały, że język ojczysty migrantów długoterminowych jest wrażliwy na zmiany w środowisku -nawet krótkotrwałe ponowne zanurzenie w L1 skutkuje zmianą strategii kontroli językowej. Badanie 3 wykazało, że krótkoterminowy trening przełączania się między językami - symulujący konkurencyjne i kooperacyjne użycia języka - również skutkuje zmianą procesów kontroli poznawczej osób dwujęzycznych. W szczególności, odkryliśmy, że konkurencyjny kontekst językowi bardziej angażuje procesy

monitorowania. Podsumowując, uzyskane wyniki wskazują na duża adaptacyjność systemu językowego, szybko zmieniającego mechanizmy poznawcze w oparciu o bieżące wymagania środowiska. Wyniki trzech przedstawionych badań poszerzają istniejącej wiedzy z zakresu psychologii, zapewniając wgląd w mechanizmy leżące u podstaw korzyści poznawczych związanych z dwujęzycznością.

1 Introduction

An ongoing debate in cognitive neuroscience is the nature of the "bilingual advantage": whether bilinguals - due to their language use - develop more efficient general-domain control processes than monolinguals (Bialystok, 2007). Broadly speaking, this advantage might result from bilinguals' constant need to control both their first language (L1) and their second language (L2) during lexical selection, as both of their languages are always activated and compete against each other (Thierry & Wu, 2007). To avoid interference from the unintended language, cognitive control mechanisms are needed to regulate the activation of both languages and allow the selection of the intended language. One possible control mechanism is the inhibition of the unintended language (D. W. Green, 1998); another is an increase in the activation level of the intended language (Costa et al., 1999; Finkbeiner et al., 2006). Regardless of the precise mechanism of language control, bilingual speakers are thought to be exposed to a greater demand on their cognitive control systems than monolingual speakers (Bialystok, 2007). This greater demand could lead to an adaptation of the cognitive control system of bilinguals and a supposed greater proficiency of bilingual speakers in cognitive abilities involved in language control (Bialy]stok, 2007). When bilinguals apply their cognitive control to non-linguistic tasks, they should show greater efficiency in cognitive mechanisms shared between tasks, resulting in a bilingual advantage of bilinguals over monolinguals (Bialystok, 2017). Many studies in recent years tried to establish this bilingual advantage and identify its precise mechanisms. While some researchers found benefits for bilingual speakers compared to monolinguals (Bialystok, 2009, 2017; Costa et al., 2008), others did not find confirmatory results (Lehtonen et al., 2018; Paap, Johnson, & Sawi, 2015, for a review of both sides see van den Noort et al., 2019). To resolve this inconsistency, researchers have suggested viewing bilingualism not as a monolithic phenomenon but instead as a collection of different language experiences, individual differences, and varying contexts (Bak, 2016; Beatty-Martínez & Titone, 2021; D. W. Green & Abutalebi, 2013; D. W. Green & Wei, 2014;

Gullifer & Titone, 2020; Kałamała et al., 2023; Navarro-Torres et al., 2021; Wodniecka, Casado, et al., 2020). Focusing on different patterns of language use and different interactional contexts should make it possible to identify underlying mechanisms that would otherwise be blurred by treating bilingualism as a categorical phenomenon.

The three investigations presented within this dissertation examine exemplary differences of regulatory, competitive, and cooperative contexts of language use - the three language contexts suggested by Beatty-Martinez and Titone (2021). The doctoral thesis is structured as follows. First, I present the theoretical background of my research and highlight exemplary consequences of changing between the language contexts suggested by Beatty-Martinez and Titone (2021). Afterward, I give an overview of each of the three investigations included in this thesis and summarize their findings. Finally, I discuss the overarching implications of this research and how it might illuminate how we describe and examine language environments and their impact on language use and cognitive control.

1.1 Describing language experience

In order to explain the often-diverging results while studying the cognitive consequences of bilingualism, different categories of bilingual language experiences have been proposed to describe forms of language interaction and their possible impact on linguistic and non-linguistic control. One influential model is the seminal Adaptive Control Hypothesis by Green and Abutalebi (2013). It suggests three interactional contexts of bilingual language use, each with specific use of languages and different resulting demands for using cognitive control abilities. The three contexts are the single language context, in which languages are confined to distinct environments; the dual language context, in which two or more languages are present within the same environment; and the dense code-switching context, in which languages are mixed within single utterances. Within each context, bilinguals must rely on different cognitive mechanisms to reliably use their language system. For instance, the dual language context requires increased use of response inhibition in order to prevent answering in the wrong language. In contrast, the code-switching context allows the free switching between languages as every speaker understands both languages and, therefore, relies to a lesser degree on response inhibition (ACH, Green & Abutalebi, 2013). Following the adaptive control hypothesis, the environment determines how bilingual speakers need to control their languages in a given situation. This, in turn, determines which language control and domain-general control processes are necessary.

In their recent review, Beatty-Martínez and Titone (2021) followed a similar line of thinking, but instead of focusing on the direct impact of the environment, they investigated how languages may influence each other. They proposed three different categories of interactional contexts that describe how the language system adapts to dynamic changes in the language environment. Within each context, languages impact each other resulting in beneficial or detrimental consequences for word production. According to Beatty-Martinez and Titone, the three possible categories of language interaction are regulation, competition, and cooperation.

Language regulation encompasses language environments and contexts "that involve variable kinds of conversational exchanges, requiring bilinguals to closely monitor and regulate the activation of both languages to suit demands in everyday life […] and manage potential between-language interference by keeping the appropriate language active while seeking new contextual cues that may signal a language change" (Beatty-Martínez & Titone, 2021, p. 5). As languages are not clearly separated from each other and speakers are not able to reliably predict the occurrence of a given language, bilinguals need to carefully regulate their language system in order to use their language successfully. The regulative context can therefore be seen as a high language entropy[1] environment. That is, the two languages are both widely spread, and speakers encounter and use them both to a similar degree (Gullifer & Titone, 2020). One such example would be L2 immersion, in which L1 bilinguals live in an L2 dominant environment and are required to adjust their L1 activation to achieve fluent L2 production. Recent studies revealed that this L2 immersion might lead to decreased L1 accessibility (Baus et al., 2013; Linck et al., 2009) and that these L2 immersed bilinguals rely more on proactive control processes such as goal maintenance or monitoring (Beatty-Martínez et al., 2020; Zhang et al., 2021). Language competition refers to a specific subset of language environments where both languages are present but separated into distinct communicative contexts. They are characterized by a low language entropy as both languages are highly compartmentalized, and speakers are able to confidently predict when they should use a given language (Gullifer & Titone, 2020). Because bilinguals in these environments do not frequently switch between languages, they show greater difficulty when required to switch within a conversation. Conversely, they tend to use more reactive language control strategies,

---

[1] Language entropy serves as an index of how diversly languages are distributed over a given environment or individual experience. High language entropy describes contexts in which two or more languages are equally likely to be engaged, while contexts with low language entropy are characterized by a single dominant language and a strong compartmentlization of languages (Gullifer & Titone, 2020).

such as inhibiting the non-target language, to use one language at the cost of another (Beatty-Martínez et al., 2020). Finally, language cooperation refers to how some environments allow the beneficial use of multiple languages within the same context. In such environments, answering in either of the languages is a valid option because all speakers use both languages; therefore, bilinguals can choose their language freely, usually based on internal factors such as lexical access, and benefit from freely mixing languages (de Bruin et al., 2018; Gollan et al., 2014; Kleinman & Gollan, 2016). Furthermore, previous studies investigated word production involving free or voluntary language choice. They found that voluntary language switching leads to reduced activation of brain areas related to language control (Blanco-Elorrieta & Pylkkänen, 2017), suggesting that voluntary language mixing requires less cognitive effort during language control. Within this context, the opportunistic selection of languages based on bottom-up processes, like lexical frequency, can lead to less effortful language use with multiple languages "cooperating" with each other instead of causing interference. Together, these three contexts form a framework to categorize the experience of different bilingual speakers according to how their languages are separated in a given context and how effortful their required language control is. It not only proposes how languages impact each other but also how languages could change over time as the result of differing language use, for example the complete suppression of a language under the competitive context or the introduction of intrasentential mixing in the cooperative context. Compared to the model proposed by Green and Abutalebi (2013), Beatty-Martinez, and Titone offer a broader classification of language experiences that is not just restricted to cognitive control but also includes effects on language use itself.

1.2 Moving from L2 immersion to L1 immersion

One difference between the language competitive and regulatory contexts is the constant need for language control in the latter. The exposure to L2 without the possibility of strictly separating the languages according to the environment (as is the case in the competitive context) requires additional control mechanisms to manage interference between the two languages. It has been suggested that this additional language control might consequently lead to hindered access to L1 in bilinguals living in L2 dominant environments and that bilingual speakers from an L2 dominant environment suffer from reduced L1 availability during speech production (Baus et al., 2013; Botezatu et al., 2020; Linck et al., 2009). Specifically, Baus et al. (2013) found that L2 immersion can reduce lexical access to L1 for low-frequency words. They tested German-Spanish bilinguals at the beginning and end of an exchange semester in Spain with a picture naming task in L1. During the second recording, after six months of L2 immersion, the exchange students showed significantly slower naming latencies for pictures with low lexical frequency names. The authors argued that these results demonstrated a reduction in L1 lexical access and that it resulted from changes in the frequency of use of their L1. However, not all studies find a clear link between L2 immersion and greater difficulty in lexical L1 access. For instance, Yilmaz and Schmid (2012) compared Turkish-Dutch bilinguals living in an L2 environment to Turkish monolinguals from an L1 environment. Both groups completed a picture naming task in L1, including stimuli with low and high lexical frequency names. The authors found no differences in picture naming performance between groups, even when accounting for lexical frequency. However, during an additional free speech task, the monolinguals showed higher diversity and sophistication in their vocabulary in the form of a greater number of individual low-frequency words than the L2 immersed bilinguals. Similar results for lexical access were found by Beatty-Martínez et al. (2020) when they compared Spanish-English bilinguals living in an L1 environment in Spain, and an L2 environment, in the United States. They also found no significant difference in the naming latencies during an L1 picture naming task. However, the tested groups differed in the use of cognitive control tested by the AX-CPT (Braver et al., 2001; Ophir et al., 2009). During this cognitive task, participants need to press a button for specific cue-stimuli combinations. By calculating whether participants are more inaccurate on wrong cue or stimuli trials, it becomes possible to determine if they rely more on proactive control (cue monitoring) or reactive control (response inhibition for wrong stimuli trials). Beatty-Martínez et al. (2020) found that L2 immersed bilinguals use more proactive control during the AX-CPT task than L1 immersed bilinguals. Given the absence of lexical access difference between L1 and L2 immersion in the presented studies, it is not clear how exactly lexical L1 access is affected by the language environment. How does L1 access change after long-term L2 immersion and short-term L1 reimmersion?

1.3 Underlying mechanisms of L1 and L2 immersion

Language regulation describes a language context in which L1 and L2 are intermixed to varying degrees so that bilingual speakers must carefully control their language system. Because languages cannot easily be separated by the environment, like in the language competition context, bilinguals must rely more closely on language control. Especially proactive control processes are used by bilingual speakers during L2 immersion, as shown by Zhang, Kang, Wu, Ma, and Guo (2015). They tested the short-term effects of language switching training on non-linguistic cognitive control with forty-eight Chinese-English bilinguals. Half of the participants were part of the experimental group and completed language training in the form of a language switching task over ten days. The AX-CPT task was applied at the beginning and end of the ten days for all participants to assess proactive and reactive control use. The authors found that the experimental group improved their performance in the AX-CPT task after language training and shifted to greater use of proactive control after language training. The control group, on the other, showed no signs of changing to either more proactive or reactive control. In a later continuation of this study, Zhang, Diaz, Guo, and Kroll (2021) compared this data to results obtained from a matched bilingual group living in an L2 dominant environment. Surprisingly they found no effect of language switching training for the L2 immersed group. In fact, the L2 immersed group showed increased use of proactive cognitive control during the experiment's first session compared to the control group. The authors concluded from these two studies that L2 immersion must encourage the use of proactive cognitive control, similar to language switching training. Therefore, it seems reasonable to assume that L2 immersed bilinguals rely heavier on proactive cognitive control compared to L1 immersed bilinguals due to the constant management of L2 interference for maintaining stable L1 accessibility. In summary, Zhang et al. (2021) demonstrated that L2 immersed bilinguals use different forms of control to meet the demands of their respective language environment.

However, even with a better understanding of how cognitive control might be guided by the environment, the exact relationship between the language environment and word production itself is still not fully understood (Section 1.2). Behavioral measures often simply form the sum of all underlying processes and require careful experimental designs in order to learn about cognitive subcomponents. On the other hand, electrophysiological data in the form of event-related potentials (ERP) can provide detailed insight into the precise timing of word production. Two ERPs established in previous literature that can inform us about lexical selection during word production (Indefrey & Levelt, 2004) are the P2 and N300 components. The naming P2 has been related to lexical access, with higher amplitudes reflecting greater retrieval difficulty (Strijkers et al., 2010, 2011, 2013). For instance, low-frequency words, which take longer to retrieve than high-frequency words, are related to higher amplitudes of the P2 component (Strijkers et al., 2010). In the same line, Strijkers et al. (2013) found an effect of language on the P2 amplitude. They tested 40 Spanish-Catalan speakers who named pictures with low- and high-frequency names. Half of the participants first named one block in L1, followed by naming in L2, while the other half named the blocks in the reverse order. The authors found a more positive P2 amplitude for L2 naming than L1 naming, but only for those participants who first named in L1. They argued that the more positive P2 amplitude during L2 naming after L1 naming resulted from additional language control mechanisms applied on the lexical level of speech production (Indefrey & Levelt, 2004). Because the language representation of L2 is weaker than L1, the use of L2 after previous L1 activation requires increased language control to resolve interference between the two languages. Branzi, Martin, Abutalebi, and Costa (2014) also associated language control applied to lexical L1 access with the P2 component. They reported that an increased frontal P2 accompanied recent L1 naming after L2 use compared to L1 naming after L1 use. The authors argued that the component was related to control processes caused by the prior use of L2. Finally, another component that was related to the ease of lexical access is the N300 component. It had previously been related to the integration of picture recognition with the name retrieval process (Wodniecka, Szewczyk, et al., 2020). In particular, the study revealed higher N300 amplitudes for naming in L1 after a block of naming in L2, which included longer naming latencies compared with naming in L1 after a block of naming in L1. The higher N300 amplitudes reflected difficulties in lexical access evoked during the previous L2 use. Together these components can give insight into the lexical selection process during word production and how it differs between bilingual speakers from the language regulation context or the language competition context.

1.4 Opportunistic language switching

Using two languages does not always have to result in interference between them. Following previous studies, Beatty-Martínez and Titone (2021) suggested that L1 and L2 can interact opportunistically without the need for additional language control processes. Within these cooperative contexts, bilinguals are able to opportunistically use the lexical availability of words within their language system and use whichever language comes first to mind, thereby increasing their word production speed. Studies on voluntary language switching found that when bilinguals switched from cued language switching to voluntary language switching, they could improve their language performance with faster naming latencies by freely switching between their languages based on internal bottom-up processes such as lexical accessibility. When bilinguals are able to use each language depending on the availability of a given word in L1 or L2, they could reduce their switch costs and decrease naming latencies overall with, in some cases, even greater speed than monolingual speakers (de Bruin et al., 2018; Gollan et al., 2014; Jevtović et al., 2020; Kleinman & Gollan, 2016). For instance, Kleinman and Gollan (2016) investigated which types of language switches are more effortful for bilingual speakers. They compared bilingual participants in a cued language switching block, two single language blocks for L1 and L2, and a bottom-up block, during which they were asked to use the same language for the same picture throughout the task but could decide the language when they first encountered the picture. They found that bilinguals were generally faster in the bottom-up task than the cued language switching task with no significant switch costs. Participants were even as fast as the single language block in the second half of the bottom-up task. The authors found that the participants showed almost no signs of difficulty in language switching because they were able to choose their preferred language for a given stimulus. Because the participants were able to choose their language for a given stimulus based on initial lexical availability, Kleinman & Gollan, 2016 argued that the choice of language for a given stimulus was already determined due to the stimulus's initial lexical availability and language control only needed to reinforce this selection process. Because language control and lexical availability aligned with each, the language switching required fewer cognitive resources. This example of cooperative language use based on lexical availability demonstrates that bilinguals can benefit from using two languages interchangeably within the same language context.

Nevertheless, it is unclear how the cooperative use of two languages affects specific language control and domain-general cognitive processes. One of the most prominent predictions comes from the adaptive control hypothesis (ACH), which emphasizes that language control processes adapt to the unique demands of each context and how each might impose different demands on different cognitive processes. One of their proposed contexts of language use is the dense code-switching context in which bilinguals routinely switch languages within a single utterance. This behavior is generally prevalent in environments dominated by bilingual speakers fluent in both their first and second language, allowing for language switches without deference to the language knowledge of the conversational partner. Specifically, the dense code-switching context encourages opportunistic planning of language use but does not increase the use of other cognitive processes, such as response inhibition or task disengagement (D. W. Green & Abutalebi, 2013). Similarly, Blanco-Elorrieta and Pylkkänen (2018) argued that not all forms of language use, specifically language switching, require increased involvement of cognitive processes. They proposed that only language contexts that externally impose language switching require the activation of cognitive processes to a relevant degree. We followed this idea in a previous study (Kałamała, Walther, et al., 2021). We developed an experimental design to artificially simulate the ACH's single and dual language contexts in an ecologically valid setting. Instead of cued language switching, we investigated naturally occurring language switches within free speech, encouraged by switching the conversational partner. We found that using L2 in both the single and the dual language context activates neural mechanisms related to response inhibition. In summary, Green and Abutalebi (2013) and Blanco-Elorrieta and Pylkkänen (2018) concluded that external constraints on language choice from the environment are major contributors to the need for language control. Bilinguals from environments with external restrictions on language use would be more likely to engage control processes to a larger degree than bilinguals from environments with cooperative language use.

2 Overview of the research program

The presented dissertation explores exemplary differences between language environments and how they affect bilingual speakers while focusing on Polish-English bilinguals and investigating both behavioral and electrophysiological measures during picture naming in L1 and L2.

I completed two experimental studies followed by three separate analyses designed to assess different types of language interactions and their consequences. Investigation 1 focused on behavioral changes to the language system due to short-term changes in the language environment in bilinguals that travel from a language regulation context to a language competition context. Investigation 2 assessed electrophysiological markers of changes in lexical L1 access assessed earlier in Investigation 1. Finally, Investigation 3 simulated changes in language environment in the laboratory settings by artificially exposing bilinguals to either a competitive or cooperative language context and assessing the impact of these contexts on participants' domain-general cognitive control.

Together, the three investigations provide data regarding the real-life impact of L2 immersion, as well as artificial environmental changes, informing us how bilinguals can vary in their language experience and how the language system can readily adapt to any given environment. Furthermore, this dissertation's experiments can help to expand our understanding of the relationship between languages and cognitive processes.

2.1 Preview of Investigation 1

Investigation 1 aimed to test if bilinguals living in an L1 environment differ in lexical L1 access from bilinguals living in an L2 environment and how changing the environment impacts lexical L1 access. In order to get a better understanding of these processes, it is necessary to remove some confounding factors affecting the previous studies. First, Yilmaz et al. (2012) compared monolinguals with bilinguals, intermixing the effects of bilingualism (knowing one or two languages) with the effects of the environment (L1 immersion against L2 immersion). Secondly, Baus et al. (2013) tested exchange students and L2 learners, while Yilmaz et al. (2012) and Beatty-Martínez et al. (2020) used long-term migrants. By combining a within-group comparison of migrants and a between-group comparison of the migrant population and a control population, the experiment tried to improve on the previous studies and establish L1 availability in the regulatory context (L2 immersion) and competitive context (L1 immersion). As L2 immersion requires extensive language control, we expected to observe greater difficulties in L1 availability compared to L1 immersion. Moving from L2 immersion to L1 immersion, on the other hand, should reduce language control and improve language availability in L1.

Sixty-four participants were tested, all Polish-English speakers with high L2 proficiency. Half of the participants were part of the control group and living in their L1 environment of Poland, while the other half consisted of long-term migrants residing in the L2 environment of the United Kingdom. All participants were tested twice: the migrant group was tested once immersed in the L2 environment and once after short-term reimmersion to the L1 environment of Poland. The control group was tested twice in the same L1 environment. Each group completed a picture-naming task in L1 with the naming latencies recorded as the experimental measure.

In contrast to our initial prediction, we found no significant difference between L2 immersed participants, and L1 immersed participants. Therefore, as we could not sufficiently establish greater difficulties for lexical L1 access during L2 immersion compared to L1 immersion, our results align with Yilmaz et al. (2012) and Beatty-Martínez et al. (2020). Still, our second prediction was confirmed as we found faster naming latencies for high-frequency words after L1 reimmersion within the migrant group, similar to Linck et al. (2009). Together these results indicate that the language system is able to compensate for long-term L2 immersion and still maintain lexical L1 access but can also quickly adapt when moving from the regulatory context to the competitive context. During this adaptation, high-frequency words are seemingly the first to benefit from this change, demonstrating the receptiveness of the language system toward environmental influences.

Reference: Casado, A., Walther, J., Wolna, A., Szewczyk, J., Sorace, A., & Wodniecka, Z. (2023). Advantages of visiting your home country: How brief reimmersion in their native country impacts migrants' native language access. Bilingualism: Language and Cognition, 1-12. doi:10.1017/S136672892300024X

2.2 Preview of Investigation 2

Investigation 2 explored neural markers of lexical access and language control under two different language environments – L2 immersion and L1 reimmersion. For this purpose, we performed analyses of electrophysiological recordings collected within the same study as Investigation 1. Instead of behavioral data, we focused on ERPs like the P2 (Strijkers et al., 2013) and N300 (Wodniecka, Szewczyk, et al., 2020) components that were previously associated with language availability. Previous studies using P2 (Branzi et al., 2014; Strijkers et al., 2013) focused only on short-term changes in the language context during a laboratory setting in the form of picture naming blocks, and it is not clear how they can be generalized to long-term immersion in a foreign language context. To explore whether bilinguals from an L2 immersion environment (language regulation) have greater difficulties accessing L1 than bilinguals from an L1 immersion environment (language competition), we focused on the analysis of the electrophysiological data of Polish-English long-term migrants living in an L2 environment before and after their return to an L1 environment. As we used established markers of the difficulty in accessing L1 words, the P2 (Strijkers et al., 2010, 2011, 2013) and N300 (Wodniecka, Szewczyk et al., 2020) components, we expected higher amplitudes for both components during L2 immersion in the migrant group compared to the L1 immersed control group. On the other hand, after L1 reimmersion, the migrants should show electrophysiological markers similar to the control group.

In total, data from 74 participants were analyzed for Investigation 2, all Polish-English speakers with high L2 proficiency. Thirty-seven of the participants were living in their L1 environment of Poland, while 37 were long-term migrants residing in the L2 environment of the United Kingdom. Identical to Investigation 1, the migrant group was tested during L2 immersion and after L1 reimmersion while also comparing their data to two sessions of the control group during L1 immersion.

We found an increased P2 amplitude in the migrant group during L2 immersion compared to the amplitude after L1 reimmersion but no overall group difference between migrants and controls. However, the N300 component was not responsive. Interestingly the observed P2 component differed in distribution from the one described in Strijkers et al. (2013) and also showed no sensitivity toward lexical frequency. We therefore argued, similar to Branzi et al. (2014), that the component could represent additional proactive language control applied during the lexical selection process for word production during L2 immersion. Previous studies (Beatty-Martínez et al., 2020; Zhang et al., 2021) already suggested that bilinguals living within a language regulatory context, L2 immersion, tend to use more proactive language control strategies in order to manage interference in their daily language use. Consequently, our results support the notion of the language system adapting to a given language environment by applying different cognitive strategies. Specifically, by using proactive language control, it is able to maintain lexical L1 access.

2.3 Preview of Investigation 3

During Investigation 3, we tested how cooperative and competitive use of language switching impacts non-linguistic cognitive processes. To answer which cognitive abilities are affected, we made use of the transfer of efficiency in shared cognitive mechanisms from language-switching training to non-linguistic cognitive tasks studies (Kałamała et al., 2021; Timmer et al., 2019; Zhang et al., 2015). The transfer of language control to domain-general cognitive control allows the selection of specific cognitive abilities and the testing of how different language environments impact their usage. We used voluntary and forced language switching to simulate cooperative and competitive language contexts in an experimental setting. Given that the competitive context is assumed to rely heavier on short-term language control than the cooperative context to correctly switch between languages (Blanco-Elorrieta & Pylkkänen, 2018; D. W. Green & Abutalebi, 2013), we hypothesized that the forced language switching group would show smaller switch cost/s in task switching after training compared to the voluntary language switching group. We expected no group difference in mixing costs, as both groups should be exposed to a similar degree to L1 and L2 use and consequently need similar global language control. Relevant event-related components will be selected for exploratory analysis.

In the final analysis, the data of 85 participants were included. They were split into two closely matched groups that completed two types of language-switching training. One group used external cues to indicate language switches, while the other could switch between languages voluntarily. By comparing the group's performance in a non-linguistic task-switching task before and after training, it was possible to assess the different impacts of competitive and cooperative language use on domain-general cognitive processes. We analyzed both the behavioral data in the form of naming latencies and the electrophysiological data as the established P3 component (Barcelo et al., 2006; Timmer et al., 2017).

The behavioral and ERP data aligned, as we found no change in switch costs but a reduction in mixing costs for the forced language switching group in the second session for both measures. In contrast to our predictions, these results demonstrate that external constraints on language use (i.e., cues during language switching) encourage greater use of cognitive global long-term mechanisms instead of transient short-term mechanisms. By trying to use both L1 and L2 in a language competition context with heavy demands on the correct language use, the forced language switching group was forced better to manage their language system on a global scale, not just on a trial-to-trial basis. The voluntary language switching group, in a language cooperation context, on the other hand, could rely on more bottom-up processes, such as lexical frequency, to guide their language. By following language availability, the participants were able to rely less on language control.

3 Investigation 1: Advantages of visiting your home country: how brief reimmersion in their native country impacts migrants' native language access

Alba Casado [1,2]*, Jonas Walther [1]*, Agata Wolna [1], Jakub Szewczyk [1,3], Antonella Sorace [4] & 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), University of Granada, Granada, Spain
3. Donders Institute for Brain, Cognition and Behaviour, Nijmegen, the Netherlands
4. School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, Edinburgh, UK

*Joint first authorship

Acknowledgments

The research was possible thanks to a National Science Centre grant (2015/18/E/HS6/00428) awarded to Z.W. 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.

Keywords: Bilingualism; L1 lexical access; L2 immersion; L1 reimmersion.

3.1 Abstract

The study explores how native language (L1) lexical access is affected by immersion in a second-language (L2) environment, and by short-term reimmersion in the L1 environment. We compared the L1 picture-naming performance of Polish-English bilinguals living in the UK (migrants) against that of bilinguals living in Poland (controls). Each group was tested twice: the migrants while in the UK (L2 immersion) and after visiting Poland (L1 reimmersion); the controls twice in their L1 environment. Contrary to our expectations, there was no main effect of group, thus suggesting that L2 immersion per se does not impact L1 lexical access. Nevertheless, migrants benefitted from L1 reimmersion by showing faster naming latencies for high-frequency words after a short visit to their home country, probably due to more opportunities to encounter these words. Overall, the study shows that the cognitive system is sensitive to the language environment by quickly adapting the activation level of lexical items.

Keywords: Bilingualism; L1 lexical access; L2 immersion; L1 reimmersion.

3.2 Introduction

When people leave their home country to live in another country, they are often required to use the second language (L2) intensively on a daily basis because most of the surrounding information, especially outside home, is presented in L2. This surrounding L2 experience is characterized by intensive use of L2 and is referred to as immersion in the L2 environment. This immersion, while being beneficial for speakers' L2, often results in them experiencing difficulty in using their native language (L1).

Previous research found that L2-immersed speakers show greater difficulties than L1-immersed speakers in producing words in L1 (picture naming: (Ammerlaan, 1996; Hulsen, 2000); verbal fluency: (Schmid & Jarvis, 2014; Schmid & Keijzer, 2009; Yagmur et al., 1999)). It was also shown that the spontaneous speech of L2-immersed speakers during L1 use is characterized by reduced vocabulary diversity (Schmid & Jarvis, 2014; Stolberg & Münch, 2010; Yilmaz & Schmid, 2012), and they frequently experience 'tip of the tongue' states (Ecke & Hall, 2013). Effects of immersion in L2 have also been observed in the processing of interface structures in L1, i.e., utterances whose meaning is determined by the combination of syntax and pragmatics, such as anaphoric pronouns (Chamorro et al., 2016; Gargiulo & van de Weijer, 2020; Tsimpli et al., 2004). The difficulty in using L1 (or, more generally, changes in processing of L1 that are experienced by speakers immersed in L2) result from an overall decrease in L1 accessibility and is known as "first-language attrition" (Köpke & Schmid, 2004). This increased difficulty in accessing or using the native language may have different origins. Some accounts have suggested that this difficulty is due to cognitive changes caused by handling two languages at the same time (Chamorro & Sorace, 2019; Sorace, 2016); others have suggested that it comes from the L1 regulation that is necessary when using L2, and/or from interference from L2 when using L1. This interference is a consequence of both of a bilingual speaker's languages being constantly co-activated (Thierry & Wu, 2007). To avoid interference from the unintended language, a control mechanism is needed to regulate the activation of both languages and allow the selection of the intended language. One possible control mechanism is inhibition of the unintended language (D. W. Green, 1998, p. 98); another is an activation increase for the intended language (Costa et al., 1999; Finkbeiner et al., 2006). The constant need to deal with interference from the unintended language reduces the efficiency of information integration and updating (Chamorro & Sorace, 2019; Sorace, 2016). Additionally, more passive mechanisms have also been proposed in which language selection is a function of the relative activation level of each language, which in turn depends on a number of factors, such as lexical frequency (Gollan et al., 2008), communicative context, the recency of language use, and the conceptual message (Blanco-Elorrieta & Caramazza, 2021).

Here, we will focus on one domain in which L2-immersed speakers may experience difficulties when using L1: accessing single words during picture naming. The aim of the current study is twofold: first, to explore the effect of long-term L2 immersion on L1 access during picture naming; second, to explore whether the consequences of L2 immersion for L1 lexical access can be reversed after short-term re-exposure to the L1 environment. In the following sections, we will review previous literature that has explored both these issues.

3.2.1 Evidence for reduced L1 access in speakers immersed in the L2 environment

Previous studies took three different approaches to studying L1 lexical access in bilinguals immersed in the L2 environment: 1. between-group comparison (bilinguals immersed in L2 environment vs. monolinguals in L1 environment); 2. between-group comparison (bilinguals immersed in L2 environment vs. bilinguals in L1 environment); 3. within-group comparison (pre-L1 reimmersion vs. post-L1 reimmersion).

3.2.1.1 Between-group comparison: bilinguals immersed in L2 environment vs. monolinguals in L1 environment

The phenomenon of reduced native language access in speakers immersed in an L2 environment has typically been explored by comparing groups of bilinguals living in an L2 environment for a relatively long time (typically migrants) against groups of monolingual speakers living in an L1 environment (Chamorro et al., 2016; Schmid & Jarvis, 2014; Sorace & Filiaci, 2006; Yagmur et al., 1999; Yilmaz & Schmid, 2012). In particular, these studies focused on some aspects of sentence processing in L1, such as anaphora resolution, lexical L1 diversity in free speech, and lexical access during L1 verbal fluency tasks. The results revealed that, compared to monolinguals, bilinguals immersed in an L2 environment had a weaker tendency to resolve anaphoric references by following the typical L1 pattern; also, their speech had lower L1 vocabulary diversity, and their semantic fluency was reduced. Still, it seems that access to the native language is not permanently lost and can be regained with L1 re-exposure.

Yilmaz and Schmid (2012) compared language accessibility in Turkish-Dutch bilinguals immersed in an L2 environment (the Netherlands) against that of monolingual Turkish speakers living in the L1 environment (Turkey). Two tasks were used: an experimentally controlled picture-naming task and a free speech task in which participants freely conversed about everyday topics. The results showed similar naming latencies for both groups during the picture-naming task, but worse performance of bilinguals in spontaneous language production (less fluent speech; in particular, reduced use of low-frequency vocabulary). The authors interpreted this finding as indicating that when L2-immersed bilinguals are required to retrieve individual words and are able to focus their attention only on this retrieval, then their performance is indistinguishable from L1-immersed monolinguals. However, in spontaneous speech, differences were observed between the two groups, possibly due to the constrained

L1 use (limited to a reduced social sphere, which decreases the use of low-frequency words) and interference from L2.

All in all, it seems that bilinguals immersed in an L2 environment experience reduced availability of low-frequency vocabulary in L1 compared with monolinguals in an L1 environment. However, contrasting these two groups of participants has some disadvantages. For instance, it is hard to distinguish between the role of knowing a second language (bilingual vs. monolingual) and the role of the linguistic environment, namely the consequences of being immersed in an L2 environment (L1 environment vs. L2 environment). This problem has been addressed by other studies directly comparing bilinguals immersed in an L2 environment against bilinguals in an L1 environment.

3.2.1.2 Between-group comparison: bilinguals immersed in an L2 environment vs. bilinguals in an L1 environment.

An alternative approach to exploring the impact of being immersed in an L2 environment involves comparing two groups of bilinguals with similar L2 proficiency but living in L1 or L2 environments. Thanks to this, we are better able to control the impact of confounding variables such as knowledge of another language and linguistic environment. This approach was used by (Linck et al., 2009), who compared two groups of English learners of Spanish: classroom students living in the USA who participated in a Spanish language course, and exchange students living for three months in Spain. The results showed that in a semantic verbal fluency task during their study-abroad experience, the exchange students produced fewer words in L1 than the classroom students. The authors interpreted these results as indicating difficulty with L1 lexical access resulting from immersion in the L2 environment. Another recent study comparing two groups of bilinguals is the one by (Botezatu et al., 2020), who compared two groups of bilinguals: native speakers of English learning Spanish or French in the US (L1 environment), and native speakers of English learning Chinese in China (L2 environment). Among other tasks, both groups of participants completed a picture-naming task in L1. During the task, the lexical frequency of the words referring to the named pictures decreased with each trial. The results showed that bilinguals immersed in the L2 environment were slower and produced more errors than bilinguals living in the L1 environment, especially for pictures corresponding to low-frequency words. Interestingly, most of the incorrect answers produced by the bilinguals immersed in the L2 environment were high-frequency substitutions for the correct low-frequency names (i.e., using the name ‘broom’ in response to ‘rake’). This last finding suggests that bilinguals immersed in an L2 environment had difficulties in accessing the correct lexical items and instead opted to use a more readily available word, albeit an incorrect one. In sum, the results of the comparison between bilinguals immersed in the L1 and L2 environments suggest that immersion in an L2 environment indeed hinders lexical retrieval in L1, and this effect might be stronger for low-frequency words.

3.2.1.3 Within-group comparison: pre-L2 immersion vs. post-L2 immersion.

Yet another approach to measuring L1 difficulties during L2 immersion is testing participants who are living in their L1 environment after being immersed in an L2 environment (e.g., spending a semester abroad). In this case, the same population is tested more than once, and factors related to individual differences due to motivation and socio-economic status or cultural heritage are controlled for. (Baus et al., 2013) tested bilinguals at two different stages: when they first arrived in an L2 environment, and after a few months of being immersed in this L2 environment. They asked German learners of Spanish to name pictures in their L1 at the beginning and end of a semester in Spain. Participants were slower to name pictures as a result of being immersed in the L2 environment. Interestingly, this slower L1 naming was more prominent for pictures with low-frequency names. The authors argued that the bilinguals reduced their use of L1 and, in particular, they used fewer low-frequency words during immersion in the L2 environment. That is, L2 immersion affects low-frequency words more than high-frequency words due to reduced L1 use during immersion in the L2 environment.

3.2.1.4 Summary

Overall, studies investigating the impact of L2 immersion on L1 availability are still scarce, but the available evidence seems to suggest that accessing and producing words in L1 becomes more challenging during immersion in an L2 environment (for contradictory results, see (Yilmaz & Schmid, 2012), who only found differences between bilinguals immersed long-term in an L2 environment and monolinguals in a free speech task and not in a controlled picture-naming experiment). Importantly, this difficulty in L1 access seems to be more pronounced for words with lower lexical frequency than those with higher lexical frequency. In the following subsection, we will review studies exploring whether the difficulties in accessing L1 could be reversed after reimmersion in the L1 environment.

3.2.2 Evidence for increased L1 access in speakers immersed in an L2 environment after reimmersion in an L1 environment

Linck et al. (2009) published the first study that explored whether L1 difficulties due to immersion in an L2 environment can be reversed after reimmersion in an L1 environment. Previously, we discussed this study in the context of comparing bilinguals immersed in an L2 environment against bilinguals in an L1 environment (see section 3.2.1.2. “Between-group comparison: bilinguals immersed in L2 environment vs. bilinguals in L1 environment”). However, these authors also tested a sub-group of bilinguals immersed in an L2 environment again after reimmersion in their L1 environment, which made it possible to test whether the effects of immersion are reversible. In particular, they re-tested some English-Spanish exchange students six months later, after they had come back to the USA, that is, after being fully reimmersed in their native language environment. The results suggested that reimmersion improved their verbal fluency in L1. The authors concluded that the difficulties experienced during immersion in the L2 environment can be overcome by returning to the native language environment. Altogether, the results of this study demonstrate that L1 lexical access becomes more difficult over the course of L2 immersion, and a short reimmersion in the L1 environment is enough to overcome the difficulties.

Chamorro et al. (2016) also studied the effects of reimmersion, but they specifically focused on sentential aspects of L1 processing rather than lexical access. They designed a cross-sectional study with three groups of adult participants: 1) Spanish monolinguals living in their L1 environment; 2) Spanish-English bilinguals living in an L2 environment for five or more years (long-term migrants); and 3) Spanish-English bilinguals who had lived in an L2 environment for five or more years and were then re-exposed for one week or more to L1 before the experiment. All the participants completed a reading task with eye-tracking in which they were asked to read and judge the naturalness of sentences in order to explore their sensitivity to pronoun mismatches. The results showed that while the groups that were recently in contact with the L1 environment had no trouble detecting pragmatically inappropriate sentence structures, the bilinguals immersed in an L2 environment by the time of testing could not successfully identify inconsistencies, measured as gaze fixations. The lack of significant differences between the performance of the monolingual group living in the L1 environment and the bilingual group living in the L2 environment who were recently reimmersed in their L1 environment suggests that recent reimmersion in the L1 environment improved their detection of pragmatically inappropriate sentence structures in comparison with the group of bilinguals living in an L2 environment. The effects of reimmersion in the L1 environment on pronoun resolution have also been addressed in a study by Gargiulo and van de Weijer (2020). They compared groups of bilingual speakers before and after reimmersion in L1, and monolinguals living in the L1 environment. A self-paced comprehension task revealed that, compared to the L1-immersed speakers, L2-immersed bilinguals had changed their preference for interpretation of null pronouns. Interestingly, short-term reimmersion in the L1 environment reversed the changes observed in bilingual speakers. However, this effect might have been task-driven as both groups, bilingual and monolingual, performed faster in the second session compared to the first session. This suggests that the improvement in the bilingual group's performance was likely due, at least in part, to training or task learning, instead of fully due to the changes in the language environment. All in all, the results of both these studies indicate that difficulties in L1 use caused by long-term immersion in an L2 environment can be overcome. As Gargiulo and van de Weijer conclude, immersion effects seem to affect language processing rather than representations, and reimmersion effects prove that the language system is flexible and easily adaptable to the requirements of the environment.

The most recent study testing bilinguals who visited their native country (L1 environment) while living in an L2 environment is a case study by (Köpke & Genevska-Hanke, 2018), who studied the spontaneous speech of a Bulgarian-German bilingual speaker in four sessions over the course of five years. Similarly to Chamorro and colleagues (2016), they focused on the use of pronouns during language production. The first and third sessions took place while the participant was immersed in the L2 environment. The second and fourth sessions were recorded in the L1 environment, where the participant had spent two weeks prior to testing. The results showed that the participant overused pronouns in L1 during immersion in the L2 environment, thus replicating the pattern of pronoun use that is typical of German, which is the participant's L2. However, a short reimmersion in the L1 environment was sufficient to return to a level of pronoun use similar to native monolingual speakers of Bulgarian. In the same vein as Chamorro et al. (2016), the authors concluded that the differences in processing anaphoric pronouns due to immersion in an L2 environment were temporary. In summary, these studies suggest that L1 re-exposure removes the constraints placed on the L1 language system during L2 immersion.

Overall, the available evidence (Chamorro et al., 2016; Köpke & Genevska-Hanke, 2018; Linck et al., 2009) seems to indicate that the L1 difficulties caused by immersion in an L2 environment can be reversed after reimmersion in the L1 environment. Still, only one of these studies (Linck et al., 2009) focused on single-word retrieval. Furthermore, it is important to note that the studies that revealed difficulty in accessing L1 words due to immersion in an L2 environment tested individuals who were L2 learners rather than highly experienced L2 users (Baus et al., 2013; Botezatu et al., 2020; Linck et al., 2009). In contrast, two out of three studies that found that L1 difficulty can be reversed after L1 reimmersion tested individuals that were long-term migrants (Chamorro et al., 2016; Köpke & Genevska-Hanke, 2018). The differences between these participant groups seem crucial because some of the observed effects on L1 could be related to the process of L2 learning (hence adapting their language system to using the two languages optimally) rather than language immersion per se.

In contrast to L2 learners, long-term L2-immersed migrant bilinguals typically use L2 rather intensively and hence have high L2 accessibility, even if they do not achieve a high level of L2 proficiency in terms of sophisticated and rich language skills in various domains. Hence, long-term migrants may experience difficulties in L1 access caused by cross-language interference from L2 to L1 (Hopp & Schmid, 2013) rather than due to changes to their language system that are related to the mere process of learning L2 (see (Bice & Kroll, 2015).

Altogether, the results from studies testing long-term migrants immersed in an L2 environment before and after reimmersion in L1 (Chamorro et al., 2016; Köpke & Genevska-Hanke, 2018) seem to align with the results of studies testing L2 learners living in their native environment before and after immersion in an L2 environment (i.e., spending a semester abroad, Baus et al., 2013; Linck et al., 2009). These studies suggest that immersion in an L2 environment interferes with L1 availability (for contradictory results see Yilmaz & Schmid, 2012), but such effects can be reversed once the bilingual comes back to their L1 environment. However, no study has yet tested whether bilinguals immersed in an L2 environment for a longer period of time (i.e., migrants) have difficulties accessing L1 compared to bilinguals in the L1 environment, and, if so, whether these difficulties could be overcome. Studying L1 lexical access in long-term migrants should give us insights into how the L2-immersed language system reacts to changes in the language environment.

3.2.2.1 Summary

All in all, from previous literature we know that L2 immersion seems to decrease the ability to access L1. However, it is not fully clear whether long-term migrants experience difficulties in L1 lexical access in comparison with bilinguals with similar L2 proficiency living in the L1 environment. Moreover, previous studies have shown that the negative impact of L2 immersion can be reversed after a short-term reimmersion in the L1 environment. Still, to date there are no studies exploring the effect of L1 reimmersion in the lexical access of long-term migrants. Therefore, the novelty of our study resides in comparing the L1 lexical access of long-term migrants during L2 immersion and after L1 reimmersion with that of a matched group of bilinguals living in the L1 environment.

3.2.3 Current study

The goal of the current study was twofold: 1) to explore whether long-term migrants (i.e., bilinguals immersed in an L2 environment) experience difficulties in L1 lexical access compared with bilinguals remaining in their L1 environment; 2) to test if L1 access in migrant bilinguals is improved after reimmersion in the L1 environment. For these purposes, two groups of Polish-English adult bilinguals were tested. We recruited a group of Polish migrants who had lived in the UK (L2 environment) for at least two years and a control group consisting of Polish-English bilinguals living in Poland (L1 environment). Each group was tested twice with an approximate between-session interval of 102 days (SD = 56). The migrant group was tested once during immersion in the L2 environment, and once after a short reimmersion in the L1 environment. The control group was also tested twice, with a similar interval between the tests, but both times this was in the L1 environment. This approach allowed us to compare the effects of immersion in an L2 environment (Botezatu et al., 2020; Chamorro et al., 2016; Yilmaz & Schmid, 2012) and assess whether short-term reimmersion in the native language environment can reverse the effects of L1 difficulty that – based on previous studies (Chamorro et al., 2016; Köpke & Genevska-Hanke, 2018) – should be observed due to long-term migration or long-term residence in the L2 environment. Thanks to the comparison between two bilingual groups of similar L2 proficiency, we should be able to disentangle the effect of L2 immersion from the effect of bilingualism per se (i.e., knowing two languages).

We counterbalanced the order of sessions between the participants: approximately half of the migrant group were first tested while immersed in the L2 environment and then again after a recent reimmersion in their L1 environment; the other half were first tested after a recent reimmersion in their L1 environment and then again when they were immersed in the L2 environment (i.e., at least 30 days after returning from Poland to the UK). Therefore, we could account for the possible confounds of the order of sessions and task training (Gargiulo & van de Weijer, 2020) by counterbalancing the order of experimental sessions.

Participants performed a battery of language-related and cognitive tasks which tested the consequences of immersion and reimmersion in many domains. In this paper we focus on one aspect of language processing, i.e., lexical access. The ease of lexical access was measured using a blocked picture-naming task. Target words corresponding to the pictures varied in lexical frequency, which allowed us to explore possible interactions between the L1 vs. L2 environments and lexical frequency. Previous research has shown not only that high-frequency words are accessed more easily than low-frequency words (Alario et al., 2004; E. Bates et al., 2003; Cuetos et al., 1999), but also that the lexical frequency of words modulates the effects of the environment. More specifically, in the L2 environment, low-frequency L1 words were particularly harder to access (Baus et al., 2013; Botezatu et al., 2020).

3.2.4 Hypotheses

We hypothesized that migrant bilinguals immersed in an L2 environment would find it more difficult to access names in L1 compared with the control bilinguals (bilinguals remaining in their L1 environment): the migrant bilinguals would be slower to name pictures in L1 than the control bilinguals. Also, we hypothesized that migrant bilinguals would find it easier to access L1 after a recent reimmersion in their native language environment, compared to while immersed in their L2 environment. This benefit should be observed as faster picture-naming latencies in L1 immediately after L1 reimmersion. Importantly, in the control bilinguals, we did not expect any changes in the L1 naming latencies between testing sessions.

Furthermore, we expected that the language environment would interact with frequency of words to be retrieved during picture naming. In terms of frequency, following Baus et al. (2013) and Botezatu et al. (2020), we expected that low-frequency L1 words would be harder to access in the L2 environment than in the L1 environment. In other words, the migrant bilinguals should be slower than the control bilinguals to name pictures that correspond to low-frequency words. Moreover, we expected that the migrant bilinguals would benefit from L1 immersion in such a way that low-frequency names would be easier to retrieve after L1 reimmersion.

3.3 Methods

3.3.1 Participants

Two groups of participants were recruited: a group of 55 Polish-English bilinguals living in the Edinburgh area, UK (migrant group), and a group of 56 Polish-English bilinguals living in Krakow, Poland (control group). All participants received monetary compensation for their time and effort. In addition, the participants in the UK were offered a Polish book as a gift. The study met the requirements and gained the approval of the Ethics Committee of Jagiellonian University Institute of Psychology concerning experimental studies with human subjects.

All the participants had learned English as a second language and used it on a daily basis (see Table 1.1). 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). The selection criteria for participating in the study were self-reported upper-intermediate English proficiency (B2) or above; accuracy of 70% or more in the General English Test; and accuracy of 60% or more in the LexTALE test. See Table 1.1 for the language abilities of both groups.

Table 1.1

Demographic information and language experience of participants.

	Migrant group (N = 32)	Control group (N = 32)	t-test
N	32 (29 female)	32 (21 female)	
Age (years)	36.16 (6.45)	29.91 (7.48)	*t*(62) = -3.58, *p* = <0.01***
SES	6.64 (1.54)	5.94 (1.63)	*t*(62) = -1.79, *p* = 0.08
Years of education	18.35 (2.58)	17.16 (2.11)	*t*(62) = -2.02, *p* = 0.05*

Length of residence in L2 environment (years)	9.66 (4.86)		-			
Length of reimmersion in L1 environment (days)	13.37 (8.18)		-			
Time delay between L1 reimmersion and recording (days)	3.06 (1.86)		-			
Self-assessed language experience	L1	L2	L1	L2	L1	L2
Self-rated proficiency	9.82 (0.47)	7.86 (0.93)	9.69 (0.70)	7.20 (1.18)	t(62) = -0.88, p = 0.38	t(62) = -2.47, p = 0.02*
Speaking	9.64 (0.79)	7.68 (1.25)	9.53 (1.08)	6.63 (1.36)	t(62) = -0.46, p = 0.65	t(62) = -3.22, p = <0.01**
Writing	9.77 (0.75)	7.42 (1.29)	9.59 (0.76)	6.63 (1.74)	t(62) = -0.92, p = 0.36	t(62) = -2.09, p = 0.04*
Listening	9.94 (0.25)	7.91 (0.86)	9.84 (0.51)	7.59 (1.10)	t(62) = -0.91, p = 0.37	t(62) = -1.27, p = 0.21
Reading	9.93 (0.25)	8.42 (0.98)	9.78 (0.79)	7.97 (1.33)	t(62) = -1.04, p = 0.30	t(62) = -1.54, p = 0.13
Percentage of daily use	40.46 (15.94)	59.25 (15.53)	81.86 (16.48)	16.75 (11.44)	t(62) = 10.21, p = <0.01***	t(62) = -12.47, p = <0.01***
Age of acquisition (years)	-	13.05 (3.72)	-	10.03 (4.37)	t(62) = -2.97, p = <0.01**	

Language switching	4.82 (2.55)		4.19 (2.29)		t(62) = -1.04, p = 0.31
Objective L2 proficiency measures					
LexTALE (mean accuracy in %)	-	77.82 (13.24)	-	73.72 (10.12)	t(62) = -1.39, p = 0.17
Cambridge test (mean accuracy in %)	-	89.88 (9.48)	-	85.00 (7.47)	t(62) = -2.28, p = 0.03*

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) number of participants and number of women in brackets, (2) age (in years), (3) socio-economic status on a 1 to 8 scale based on Adler, Epel, Castellazzo, and Ickovics (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 between the return from the L1-environment and the experimental recording. The second part of the table summarizes 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 the age of acquisition 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.

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

For the migrant group, we recruited Polish native speakers who had lived in the UK for a minimum of two years. The migrant group was tested twice: 1) after at least 30 days fully immersed in the L2 environment, that is, without leaving the UK, which we referred to as the “during L2 immersion” session; 2) after reimmersion in the L1 environment, less than 7 days after returning from Poland, which we referred to as the “after L1 reimmersion” session. The order of the first session was counterbalanced (half of the participants performed the “during the L2 immersion session” first; half of the participants performed the “after the L1 reimmersion session” first). See Figure 1.1. From the initial sample, we excluded nine participants who did not complete the two sessions; another five were excluded because they did not follow the established time limit for each session. Additionally, four more participants were excluded due to technical problems during the recording of the responses. All of the remaining participants reported that they only used Polish when in contact with friends and family in Poland, never English or other languages. The final sample included 37 participants.

For the control group, native speakers of Polish with high English proficiency were recruited. Pre-selection criteria allowed only participants who had spent the last 30 days in Poland before each session. Similarly to the migrant group, the control group was tested in two sessions. In contrast to the migrant group, Context was a dummy variable in the control group, so we referred to these sessions as X Context and Y Context. See Figure 1.1. From the initial sample of 56 participants, we selected a subset of 37 participants such that they matched the migrant group as closely as possible on a set of critical measures: chronological age, age of L2 acquisition, socio-economic status (SES), language proficiency (the combined score of the LexTALE and Cambridge tasks) and self-assessed language switching behavior. The matching procedure was carried out using a brute force algorithm that we also used for a similar purpose in a previous study (Marecka et al., 2020). The groups had means within 1 SD of each matched variable (see Table 1.1 for the observed similarities and differences between the groups). To account for any remaining differences between the groups, we statistically adjusted for age and L2 age of acquisition in all analyses comparing the groups.

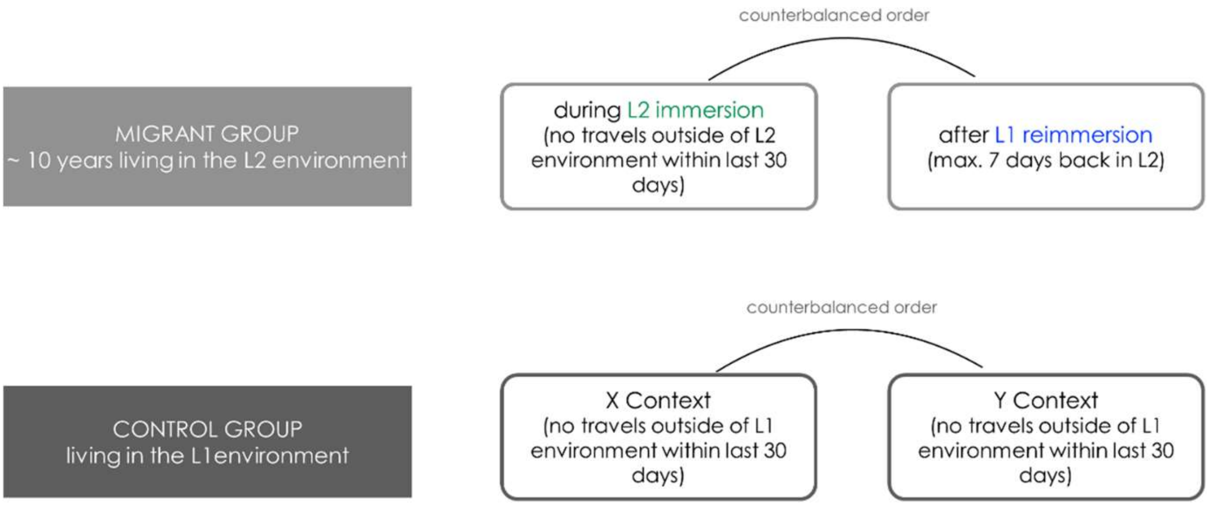
3.3.2 Task and procedure

The average time interval between sessions was 102 days (*SD* = 56; range: 30–260). The order of the sessions was counterbalanced for both groups between participants by randomly assigning each participant to their first experimental session (migrant group – 19 participants started with the after L1 reimmersion session; control group – 18 participants started with the X Context session).

3.3.2.1 Materials

We selected two different sets of pictures that were matched on a number of lexical characteristics in order to have non-repeated pictures for the two picture-naming sessions. The stimuli in the picture-naming task consisted of 216 colored images from the Cross-Linguistic Lexical Tasks database (Haman et al., 2017). We divided all the pictures into four subsets and created two different versions with different orders of presentation for each subset. The subsets of pictures were balanced with respect to name agreement, lexical frequency (based on (Mandera et al., 2015), and mean length in phonemes. Moreover, each subset contained a comparable number of images from different semantic categories. The order in which the subsets were presented was counterbalanced across participants. To avoid training effects and other confounds related to item repetition, all items were presented only once (Mitchell & Brown, 1988).

Figure 1.1



Note. Representation of the testing sessions for the migrant and control group.

3.3.2.2 Procedure

In the picture-naming task, the pictures were displayed in the center of a computer screen on a black background using DMDX (Forster & Forster, 2003). Each trial was preceded by a black screen presented for 1000 ms, followed by a fixation cross that appeared in the screen's center for 1000 ms. A picture was then shown in the center of the screen until the participant responded or until the timeout was reached (3000 ms). The participants were instructed to name pictures aloud in their native language as quickly and accurately as possible. Vocal responses were recorded as audio files using DMDX. 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.

3.3.3 Analysis

3.3.3.1 Naming latencies

The naming latencies were determined from the audio files using the Chronset online tool (Roux et al., 2017). Practice trials were not included in the analysis. Responses with naming latencies below 300 ms and trials with inaccurate naming or timeout (3000 ms) were removed from the data. In total, 8.68% of the data was excluded. Due to the right-skewed distribution of naming latencies, they were transformed using reciprocal transformation (-1000/naming latency).

3.3.3.2 Statistical analysis

We used the lme4 package (D. Bates et al., 2015, Version 1.1-23) in R (R Development Core Team, 2020, Version 4.0.2) to calculate the linear mixed-effects models.

We first fitted a general model that included both groups of participants. It also included participants and pictures as crossed random effects. As fixed effects, the model included Group (Migrant, Control) and Context (L1/X Context, L2/Y Context). As item-related fixed effects, the model included Word-lexical frequency (i.e., target name's lexical frequency based on Mandera et al., 2015) and the Trial number. We also included two participant-related fixed effects in the model: Age, and Age of L2 acquisition. Finally, the model included the interactions between Group and Context, and the interactions between Group, Context, and Word-lexical frequency. Before running the analyses, all categorical predictors were deviation coded using the sum contrast (Group: Control group = -0.5, Migrant group = 0.5; Context:

L1/X-Context = -0.5, L2/Y-Context = 0.5). Trial number was log transformed. The continuous predictors were centered and standardized (Age and Age of L2 acquisition). We used the so-called maximum random-effects structure (Barr et al., 2013): by-Picture random intercept and random slopes for Group, Context, Age, Age of L2 acquisition, Trial number, and the interaction between Group and Context; by-Participant random intercept with random slopes for Context, Word-lexical frequency, and Trial number; and interactions between all slopes and intercepts.

We fitted the maximal model first. If it did not converge, we first removed correlations between random effects; in the next step, the random effects with the smallest unique variance were removed, following the recommendation of Bates, Kliegl, Vasishth, and Baayen (2018). Absolute t-values greater than two were considered significant. The final model was: lmer (inverted_RT ~ Group + Context + Word_Lexical_Freq + Group : Context + Group: Context : Word_Lexical_Freq + Group : Word_Lexical_Freq + Age + AoA_L2 + Trial.number + (1 + Context| Participant) + (1 + Context| Item), data).

3.4 Results

3.4.1 General Model

The results revealed a significant interaction between Group, Context and Word-lexical frequency (see Table 1.3): there were faster naming latencies after the L1 reimmersion compared to the during L2 immersion, but only for words with higher frequencies (See Figure 1.2).

Table 1.2

Summary of the raw behavioural data of L1 picture naming.

Behavioural measures	Group	During L2 immersion / X-Context	After L1 reimmersion / Y-Context	t-test (Context comparison)
Naming latencies (ms)	Migrant group	973 (311)	934.76 (300)	*t*(72)= -1.27; *p* = 0.21
	Control group	1001 (313)	1019.81 (317)	*t*(72)= -0.02; *p* = 0.99
	t-test (Group comparison)	*t*(72)= 0.39; *p* = 0.69	*t*(72)= 1.29; *p* = 0.20	
Accuracy	Migrant group	0.97 (0.17)	0.97 (0.17)	*t*(72)= -0.28; *p* = 0.78
	Control group	0.96 (0.21)	0.94 (0.25)	*t*(72)= -1.14; *p* = 0.26
	t-test (Group comparison)	*t*(72)= -1.63; *p* = 0.11	*t*(72)= -1.88; *p* = 0.06	

Note. The table gives the raw behavioural measures of the two groups in the L1 picture naming task. T-tests compared between the migrant and the control group within each measure and language or context. Standard deviations are given in parenthesis.

Table 1.3

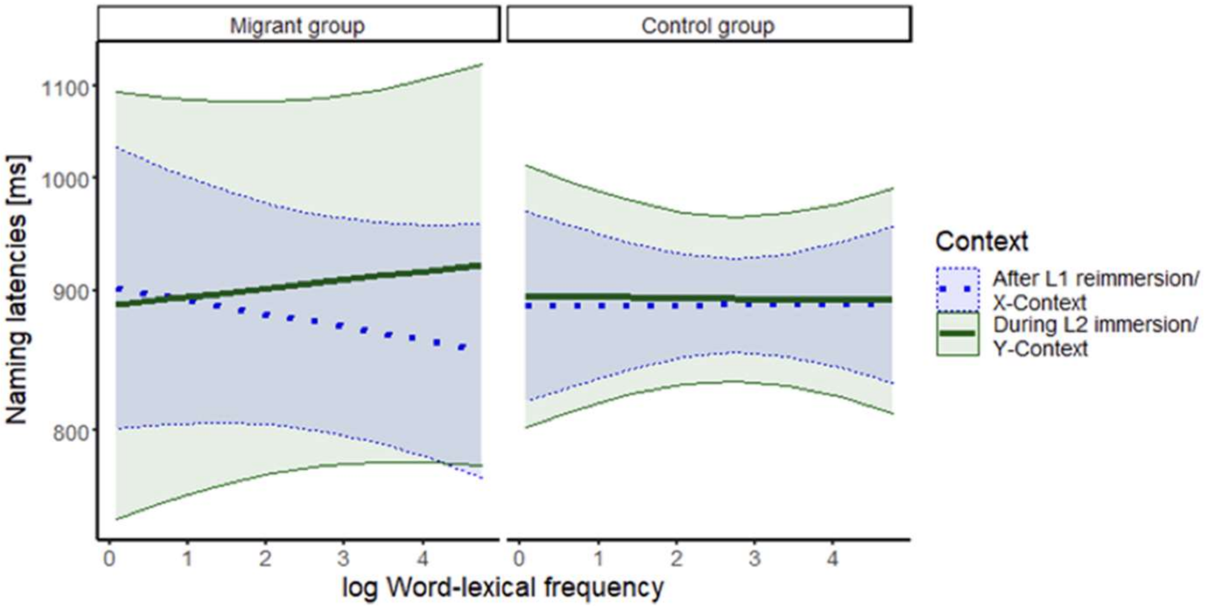
Fixed effects for the LME model for naming latencies of the general model.

Effect	Estimate	SE	t	by-Picture SD	by-Participant SD
Intercept	-1.14	0.03	-38.00***	0.15	0.18
Group	0.00	0.05	-0.02		
Context	0.02	0.04	0.65	-0.05	-0.11
Word-lexical frequency	0.00	0.01	-0.08	-	
Age	-0.03	0.02	-1.30		-
Age of acquisition	0.03	0.02	1.15		-
log (Trial number)	0.00	0.01	0.78		
Group:Context	0.03	0.07	0.48		-
Group:Word-lexical frequency	0.00	0.00	-0.28	-	-
Control Group:Context:Word-lexical frequency	0.00	0.01	-0.11	-	-
Mig. Group:Context:Word-lexical frequency	0.01	0.01	1.94'	-	-

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

Figure 1.2

Interaction of Group, Context and Word-lexical frequency in General Model



Note. Marginal effects of the final LME model for the interaction between Group, Context and Word-lexical frequency. The straight line corresponds to the L2 immersion or Y-Context; the dotted line corresponds to after L1 reimmersion or the X-Context. The error ribbon represents 95% confidence intervals.

3.5 Discussion

The present study tested how L1 lexical access is affected by being immersed in an L2 environment for a relatively long time, as well as the consequences of short-term reimmersion in the L1 environment. To this aim, we compared the picture-naming performance in the L1 (Polish) of Polish-English bilinguals who had lived in the UK for at least two years (migrants) against the picture-naming performance in the L1 of Polish-English bilinguals living in Poland (control). Each group was tested twice. Participants from the migrant group were tested while in the UK (L2 immersion) and after a short visit to their native language environment (L1 reimmersion in Poland). Participants from the control group were tested both times in their native language environment.

We formulated two main hypotheses: 1) long-term immersion in an L2 environment results in reduced lexical access in L1. Consequently, compared to the control group, which is immersed in the L1 environment, we should observe slower naming latencies in the migrant group while immersed in the L2 environment. Moreover, 2) we hypothesized that difficulties resulting from being immersed in an L2 environment should diminish after a short reimmersion in the L1 environment. That is, we should observe that the migrant group's naming latencies during their immersion in an L2 environment should be slower compared with after reimmersion in the L1 environment. Additionally, we expected interactions between the language environment and the lexical frequency of the words corresponding to the pictures to be named, because previous literature reported impact of word frequency on the ease of lexical access. In brief, we expected that lower word frequency would add to retrieval difficulty under more difficult circumstances, i.e., in the L2 environment. Below, we discuss the findings in relation to each of the formulated hypotheses.

3.5.1 Group comparison

In contrast to our main hypothesis but in line with Yilmaz and Schmid (2012), there was no main effect of group in the main analysis, nor in the t-test comparison of the raw naming latencies (see Table 1.2); this most likely indicates that, overall, there were no differences in naming latencies between the control and migrant groups. This suggests that, despite being immersed in the L2 environment for an extended period of time, the migrant bilinguals did not have difficulties accessing their L1 – at least in this highly controlled picture-naming task. The lack of group differences in our study contrasts with previous studies that found reduced L1 lexical access for bilinguals immersed in an L2 environment (Baus et al., 2013; Botezatu et al., 2020; Linck et al., 2009). A possible explanation for the absence of a similar effect in our study could be related to the different lengths and natures of the tested groups' time spent abroad. That is, the participants in our study were migrants who had already spent several years immersed in an L2 environment (range = 2 to 24 years; mean = 9.61; sd = 4.43). In contrast, previous research (Baus et al., 2013; Botezatu et al., 2020; Linck et al., 2009) tested bilinguals that were actively learning L2 at the time of testing and had spent only a few months abroad. Because L2 knowledge increases during L2 immersion, it is possible that changes in L1 performance may occur in L2 learners who are immersed in the L2 environment. Bice and Kroll (2015) showed that when bilinguals are acquiring an L2, they adapt their language system to accommodate the new language. This adaptation of the language system may rely on L1 inhibition, and this L1 inhibition may trigger difficulties in L1 access for some time. This difficulty is, however, desirable as it reduces language competition (Bogulski et al., 2019). In other words, during L2 immersion, L2 learners constantly inhibit their L1 to facilitate L2 learning. It is possible that the differences between L2 learners in the L1 environment and L2 learners in the L2 environment (Botezatu et al., 2020; Linck et al., 2009) boil down to the fact that these environments offer a different number of opportunities to learn L2. In contrast, long-term migrant bilinguals do not actively inhibit their L1 but keep it available to use when appropriate. Alternatively, it may be the case that long-term immersed bilinguals develop a very efficient inhibitory mechanism, which allows them to control cross-language competition in such a way that the consequences of L1 inhibition are not observable because of its efficient recovery (Jacobs et al., 2016).

Support for the hypothesis that immersion only affects L1 lexical access in L2 learners but not in long-term migrant bilinguals can be provided by the study by Yilmaz and Schmid (2012). Similarly to the current study, Yilmaz and Schmid tested long-term migrants immersed in an L2 environment for at least ten years. They compared the migrant group with monolinguals living in an L1 environment and found no differences in L1 lexical access in a simple picture-naming task. Accordingly, these authors argued that L1 lexical representations can remain intact despite extended immersion in an L2 environment. In line with Yilmaz and Schmid's (2012) interpretation, it seems reasonable to assume that the lexical representations of our migrant population also remain intact, even though they are immersed in their L2 environment, and that is why they access L1 effortlessly. However, while the absence of group differences in L1 access in Yilmaz and Schmid's study (which compared monolinguals and bilinguals) could have been confounded by the bilingualism factor, this was controlled for in the current study because we compared two groups of bilinguals rather than bilinguals with monolinguals. Moreover, the only difference between the groups was that, at some point, some of them migrated to the UK, where the primary language is English. All participants included in our study were raised in a similar environment (i.e., native country – Poland) and were matched on L2 proficiency. Therefore, we were able to disentangle the effect of the language environment from the effect of bilingualism (i.e., knowing more than one language). The absence of any difference in lexical access in the compared groups suggests that in long-term migrants the effects of the environment are indistinguishable from the mere fact of knowing more than one language; it also suggests that, in the migrant population, immersion in an L2 environment does not necessarily lead to a detriment in accessing L1 (assessed via picture-naming latencies). As such, we argue that the reduced L1 access found in previous studies that tested bilinguals immersed in an L2 environment might have been related to the relatively early stages of the L2 learning process and the intensity of L2 learning in the L2 environment (Baus et al., 2013; Botezatu et al., 2020; Linck et al., 2009).

Another possibility that could explain the similarity in the naming latencies between the long-term migrants and the control group is that our immersed participants maintained close contact with their native country. In particular, they were connected to the Polish community living in the UK and were mostly recruited via Polish social media. In a previous study, Hulsen et al. (2000) found that active contact with the native country influences the L1 performance of long-term migrants. These researchers tested Dutch-English bilinguals living in New Zealand for about 36 years and compared their performance in a picture-naming task with that of a Dutch monolingual group living in the Netherlands. The results showed similar naming latencies between groups. Moreover, in the migrant group, there was a partial correlation between naming latencies and how often these migrants maintained contact in Dutch (L1) with people living in the Netherlands. In particular, the migrants who had more extensive contact with their home country exhibited faster naming latencies compared with the migrants with less contact. Notably, we are not able to make a correlation because the current study lacks quantitative information about the migrant group members' contact with their home country. However, the participants were recruited using a snowball strategy and through various Polish communities and media in the UK. We therefore targeted people who maintained contact with the Polish diaspora in the UK but had obtained relatively high levels of proficiency in the majority language, namely English. Moreover, given the fact that we required the immersed participants to travel to Poland, we presumed that they maintained frequent contact with their home country. This frequent contact might have contributed to the increased L1 performance of the long-term migrants and consequently to similar naming latencies as the control group.

Altogether, it seems that our results provide initial evidence that bilinguals who are also migrants can have comparable lexical access in their native language to bilinguals who live in their L1 environment; however, at this point we cannot exclude other alternative explanations (i.e., that the lack of significant group differences is related to limited power due to the sample size of ~40 participants per group).

3.5.2 Effects of reimmersion

Our second hypothesis is related to the impact of short reimmersion in the L1 environment. Based on the previous findings of Linck et al., (2009), we expected that a short reimmersion in the native language context (a short visit to the native country) would improve the L1 lexical access of speakers usually immersed in an L2 environment (migrant bilinguals). The results showed that lexical access in L1 indeed benefitted from L1 reimmersion, but only for high-frequency words. In other words, bilingual migrants were faster to retrieve higher-frequency L1 words after a visit to their home country than while residing in the L2 environment.

Although the results are opposite to what we initially expected, they actually seem to complement previous studies which showed that L2 learners immersed in an L2 environment were slower to access words with lower lexical word frequencies (Baus et al., 2013; Botezatu et al., 2020). Our design is the reverse of the design used by Baus et al. (2013), who first assessed the participants in their L1 environment and then re-tested them after six months immersed in the L2 environment. We, on the other hand, assessed the participants after at least 2 years of immersion in the L2 environment, and also when they returned from their L1 environment (as a reminder, the order of the L2 environment session and the L1 environment session was counterbalanced). At first sight, our results seem to contradict Baus et al. (2013), who found slower naming latencies for words with lower lexical frequency when the bilinguals were tested in their L2 environment (at the end of the semester) compared to when tested in their L1 environment (at the beginning of the semester). In contrast, we found that bilinguals reimmersed in their L1 environment were faster in naming words with higher lexical frequencies compared to bilinguals immersed in an L2 environment, which is exactly the opposite pattern. A possible explanation for the differences in results between Baus et al. and the current experiment could be related to how often L1 words are likely to be used in the language environment of a bilingual speaker. Baus et al. argued that the lexical frequency effect observed in their study was caused by the relatively infrequent use of low-frequency words in their mental lexicon, compared with high-frequency words (Gollan et al., 2008). Because of the relatively infrequent use of L1 words for an extended period (i.e., while immersed in an L2 environment), the activation of L1 words is reduced. Therefore, low-frequency words (having lower activation) are the first to show noticeably reduced lexical access during L1 word production. Following the hypothesis of Baus et al., we should observe that, compared with high-frequency words, low-frequency words benefit more from L1 reimmersion due to their weaker lexical representation. Contrary to these expectations, we observed that, after reimmersion in L1, lexical access was facilitated for words of higher frequency. This seemingly contradictory finding seems, however, to make sense if we assume that the difference in the difficulty of lexical access between low- and high-frequency words is not necessarily a function of their baseline activation strength; instead, it depends on how often these words are encountered and used in each environment. That is, in the L2 environment the activation threshold of L1 words will be similarly low for high- and low-frequency words due to their reduced use in the L2 environment in general. However, L1 use usually increases during brief reimmersion in the L1 environment, therefore it is very likely that high-frequency words are encountered more often than low-frequency words. That being the case, high-frequency words benefit more than low-frequency words during L1 reimmersion because they are encountered and used more often. Therefore, the ease of lexical access observed during reimmersion in the L1 environment could be at least partially determined by the actual encountering and use of L1 words in the native language environment. This explanation could also account for the results of Baus et al.: after immersion in the L2 environment, low-frequency words trigger a higher lexical access cost because they are less likely to be used. In comparison, high-frequency L1 words might still be occasionally encountered, therefore they may be active in the language system for longer. Altogether, our findings, as well as those of Baus et al., appear to highlight the close relationship between changes in the language environment and the role of lexical frequency. By moving into a different language environment, bilinguals change their day-to-day use and exposure to L1 and L2, which leads to changes in lexical access of L1. In other words, a change in language environment increases sensitivity to word frequency. It follows that if the control bilinguals were tested in the L2 environment, we should observe frequency effects similar to what Baus et al. found, but this hypothesis requires further testing.

Following this perspective, Beatty-Martínez et al. (2020) tested how the language environment interacts with participants' patterns of cognitive control; they also explored the role of the environment in relation to the lexical frequency of words. Given the focus of our current study, we focus only on describing the interaction between the language environment and the lexical frequency effects observed by Beatty-Martinez et al., who tested Spanish-English bilinguals from three different environments: Granada (Spain) as an L1-dominant environment; Pennsylvania (United States) as an L2-dominant environment; and Puerto Rico (United States) as a language switching environment. All participants completed picture-naming tasks in L1 and L2, including items with varied lexical frequency. Similarly to the current study, no significant differences between the naming latencies of participants from the three different language environments were found. However, the results revealed that bilinguals from the L1-dominant environment showed a greater frequency effect (i.e., faster naming latencies for high-frequency words vs. low-frequency words) in L1 than in L2. In contrast, bilinguals from the L2-dominant environment demonstrated more similar frequency effects between L1 and L2. Moreover, Beatty-Martinez et al. found a correspondence between language environment and patterns of language control that could explain the different patterns of lexical frequency. Bilinguals in the L1-dominant environment tended to apply more reactive control, therefore they reacted more quickly to more-salient stimuli, i.e., high-frequency words. In contrast, bilinguals in the L2-dominant environment tended to apply more proactive control, which favored the suppression of competing items from both L1 and L2. This greater control of competing items allowed them to retrieve low-frequency items in L1 more quickly, thus reducing the frequency effects. Altogether, Beatty-Martinez et al.'s (2020) results align with those of the current study in the sense that when bilinguals are immersed in an L1-dominant environment, they experience a higher lexical frequency effect in L1 compared to when they are immersed in a L2-dominant environment. Overall, the results seem to suggest that the language system adapts to the linguistic environment, although the clear nature of this adaptation remains to be explained in future studies.

Another way to explain the interaction between lexical frequency and migrants' language environment would be that immersion in the L2 environment reduces or blocks the lexical frequency effect (i.e., high-frequency L1 words are more available than low-frequency words). As can be observed in Figure 1.2, during L2 immersion the lexical frequency does not predict the migrants' naming latencies, relative to after L1 reimmersion. It may be the case that L2 immersion blocks the frequency effect as it may become counterproductive to maintain more available high-frequency L1 words in the L2 environment because this may create interference when using the L2. However, as soon as the migrants are reimmersed in their L1 environment, the frequency effect emerges. This apparent pattern could tentatively be explained by a decrease in the frequency of using L1 during L2 immersion. That is, the decrease in L1 use would reduce the unique lexical frequency of L1 items, including high-frequency L1 items (in line with the frequency-lag hypothesis of Gollan et al., 2011). In the same line, this could also be explained as global inhibition applied to L1 during L2 immersion, thus reducing the availability of all L1 items, including high-frequency words (Van Assche et al., 2013). A slightly different explanation for the increase in naming latencies for high-frequency words during L2 immersion could be that local inhibition was applied to high-frequency L1 items, which would interfere with the use of translation equivalents in L2 in the L2 environment (in line with the local inhibition proposal by Sandoval et al., 2010).

In order to explore these possibilities, we performed a further analysis for the migrant group in which we included as covariates the years spent living in the L2 environment, and the days of reimmersion in the L1 environment (analysis presented in Appendix 8.1.2). The results showed that the length of immersion/reimmersion did not affect the main effect of context, thus indicating that this effect was most likely due to global L1 inhibition in the L2 environment (Linck et al., 2009), which is decreased after reimmersion in the L1 environment. However, further research is needed to confirm this explanation.

All in all, the language system is sensitive to the language environment and can benefit from language exposure by quickly adapting the activation level of lexical items. That is, the language system seems to be able to readily adapt to changes in the language context, thus demonstrating that long-term immersion in an L2 environment does not have a negative impact on L1 lexical access and that L1 access can improve after reimmersion in an L1 environment. The increased speed of L1 naming after reimmersion in an L1 environment extends the existing literature on the reversibility of pragmatic processing of meaning (Chamorro et al., 2016) or spontaneous speech (Köpke & Genevska-Hanke, 2018). As far as we know, these are the first results that show that even brief reimmersion in the L1 environment improves bilinguals' lexical access to their native language for those who otherwise remain immersed in an L2 environment.

3.5.3 Summary

All things considered, our data demonstrate that long-term migrants, despite being immersed in the L2 environment, maintain an ability to access L1 that is comparable to that of bilinguals living in the L1 environment. Moreover, long-term migrants can still benefit from short-term reimmersion in the L1 environment, as shown by improved access to high-frequency L1 words.

3.5.4 Limitations and future directions

The main limitation of the present study is that the effect of L1 reimmersion was not tested during the period of L1 reimmersion but up to 7 days after migrants' return to the L2 environment. This time gap, in which the participants were already exposed to the L2 environment, might have obscured some of the reimmersion effects. Additionally, despite our great efforts to match the groups, some characteristics of the migrants and controls still differed, such as L2 proficiency level, age of L2 acquisition, and chronological age. One possibility for future research would be to compare our migrant population with long-term migrants who do not maintain close contact with their L1 environment and therefore may experience different changes in their L1 accessibility.

3.6 Conclusion

To conclude, the presented study demonstrates that L2 immersion does not necessarily have detrimental effects on native lexical access, as observed in a simple picture-naming task. Instead, migrants living in an L2 environment can demonstrate a similar ease of lexical access as bilinguals living in an L1 environment. Additionally, we found that the language system is capable of quickly adapting to changes in the language environment. This was corroborated by the observed beneficial effects of a short-term visit to a home country for easiness of access to words of higher lexical frequency. This change seems to be driven predominantly by opportunities to use particular words, with more prevalent words having more impact than those encountered less often. Future research will be essential to determine the specific mechanisms underlying these adaptation processes and whether these are similar in bilinguals detached from their L1 environment. Additionally, it is relevant to explore how the language environment modulates lexical access in the second language.

Competing interests: The authors declare none

Authors' note: data available on
https://osf.io/tdxsr/?view_only=a146a3de536d4b92a70a46b3a7240f80

4 Investigation 2: Long- and short-term adaptation of the bilingual's language system to different language environments: an electrophysiological investigation

Alba Casado [1,2]*, Jonas Walther [1]*, Agata Wolna [1], Jakub Szewczyk [1,3], Antonella Sorace [4] & 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), University of Granada, Granada, Spain
3. Donders Institute for Brain, Cognition and Behaviour, Nijmegen, the Netherlands
4. School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, Edinburgh, UK

*Joint first authorship

Acknowledgments

The research was possible thanks to a National Science Centre grant (2015/18/E/HS6/00428) awarded to Z.W. 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.

4.1 Abstract

Does a long-term stay in a foreign language country affect our ability to retrieve words in our native language? And if so, are the effects reversible? The present study explored the neural correlates of lexical access in the native language and their dynamics due to a change of language environment: during long-term immersion in a foreign language (L2) environment and after reimmersion in a native language (L1) environment. We tested Polish-English migrants living in the UK (L2 environment) for about ten years and Polish-English controls living in Poland (L1 environment). All participants performed a 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. We focused on exploring two event-related components previously associated with the ease of lexical access: P2 and N300. We found that the language environment modulated the P2 component in migrants, such that higher amplitudes were evoked while naming pictures during immersion in the L2 environment compared to after the short reimmersion in the L1 environment. However, no overall differences were found when comparing the two groups. These results indicate that immersion in different language environments induces modulations in the neural response associated with word retrieval. Moreover, the modulations of P2 amplitudes related to changes in language environment seem to index higher proactive control applied during L2 immersion and its reduction after a short-term L1 immersion.

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

4.2 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' 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., 2020). These difficulties using L1 are usually discussed in the literature under "first-language attrition," and attributed to an overall decrease in L1 access. Alternative explanations could be fundamental changes in the lexicon or general language knowledge due to L2 immersion (Köpke & Schmid, 2004; Schmid, 2004, 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 of language skills from before immersion (Chamorro et al., 2016; Köpke & Genevska-Hanke, 2018; Linck et al., 2009). However, not all studies report a behavioral effect of long-term L2 immersion on L1 access (Beatty-Martínez et al., 2020; Casado et al., 2023; Yilmaz & Schmid, 2012). It is therefore of interest to explore whether the absence of behavioral effects can still be accompanied by signs of lexical access difficulty on the neural level. Notably, none of the previous studies have employed neurophysiological markers like electrophysiological indices to track changes in language processing that occur at the brain level due to language (re)immersion. The electrophysiological (EEG) technique is a tool with a precise time resolution allowing one to explore brain reactions to stimuli in a very precise time scale. As a result, event-related potentials (ERPs) can provide a measure of the processing of stimuli even when there is no behavioral change (Bice & Kroll, 2015; McLaughlin et al., 2004). It is, therefore, of interest to explore whether electrophysiological markers of lexical access reflect the difficulty in accessing the native language in migrants and whether the same markers are also sensitive to short-term reimmersion to the L1. Consequently, the goal of this study was to examine the electrophysiological response in the picture naming of Polish-English migrants in a Polish and an English environment. Below, we briefly review the differing results of studies on how L1 lexical access can be affected by the language environment.

4.2.1 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, 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 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 re-immersion to an L1 environment, there seems to be a difference in lexical L1 access (Baus et al., 2013; Botezatu et al., 2021; Linck et al., 2009). Our own behavioral data (Casado 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 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 re-immersed 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 better faster naming latencies during a blocked picture-naming task in L1 (Casado et al., 2023). In particular, in the latter study, we found that Polish-English bilinguals living in the UK showed faster naming latencies of high- vs low-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 L1, the speakers primarily encountered high-frequency words, which led to facilitation of lexical access to these words observed even a few days after their return to the L2 environment.

Altogether, based on the behavioral evidence, we can distinguish two different features 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 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. 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 behavioral tasks, it is unknown whether we could detect more subtle differences on the neural level, especially in comparison with bilinguals in the L1 environment.

4.2.2 Electrophysiological correlates of lexical access

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

The first component is the production P2, characterized by a positive activity peaking around 200 ms in central[2] electrodes (Baus et al., 2020; Strijkers et al., 2010, 2011, 2013). Larger 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 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 amplitudes of the P2 component (Costa, Strijkers, et al., 2009). Additionally, L2 words, which in unbalanced bilinguals take longer to retrieve than L1 words, were associated with higher amplitudes of the P2 component (Strijkers et al., 2013). Altogether, previous studies showed that the naming P2 component reflects difficulties in lexical access during word production.

Another component previously associated with the difficulty of retrieving words during picture naming is the N300. It is characterized by a negative activity peaking around 300 ms in central

---

[2] Previous studies characterised the P2 component at different locations: parietal-occipital (Costa et al., 2009) with a nose reference, posterior with a wide distribution (Strijkers et al., 2010, nose reference) fronto-central (Strijkers et al., 2011, left mastoid reference). Moreover, with an averaged mastoid reference 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. For the purpose of this study, we selected Strijkers et al (2013) as the template of our P2 characterization as it matches our choice of reference electrode and also investigates the effects of language.

electrodes. Previously the component has been associated with the recognition of pictures with smaller 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 of 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 (Wodniecka, Szewczyk, et al., 2020). More specifically, the increased N300 amplitude was 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 N300 amplitudes seem to reflect difficulties retrieving the lexical information associated with names of corresponding pictures.

4.2.3 Current study

In the current study, we explored electrophysiological correlates of lexical access difficulty in the native language 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 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 (a visit to their home country). 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 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: 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 system, also in comparison with control bilinguals, due to changes in the language environment.

We made three following 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 be accompanied by a more 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 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 and N300 amplitudes should also be modulated by the lexical frequency of the words; in particular, low-frequency words should evoke higher amplitudes of the P2 and N300 compared with high-frequency words. Still, to date no previous ERP study has explored this interaction, therefore we may find a different pattern.

4.3 Methods

4.3.1 Participants

We analyzed data taken from our large-scale study in which we tested the consequences of immersion and reimmersion in many language domains by collecting the data from a battery of language-related tasks. In that study, we collected a sample of 55 Polish-English bilinguals living in the United Kingdom (migrant group) and 56 Polish-English bilinguals residing in Poland (control group). In the present study, we focus on lexical access, assessed through the EEG response, 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 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, less than 7 days after returning from Poland. 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 behavioral responses. The final sample included in the analyses consisted of 32 participants[3].

For the control group, we recruited native speakers of Polish with high English proficiency. Moreover, pre-selection criteria only allowed those participants who spent the last 30 days in Poland before each session. Like the migrant group, the control group completed two

---

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

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. 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 task errors. The remaining 32 participants[4] were compared to the migrant group.

---

[4] Out of the 32 control group participants whose data were analysed for the purpose of the present study, only twenty-five control group participants were included both in this study and in the behavioural analysis presented in Casado et al. (2023), given that only these data sets had both high-quality behavioural and electrophysiological data.

Table 2.1

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)		29.84 (7.35)		t(62) = -4.19, p = <0.01***	
SES (1 - 10)	6.64 (1.54)		6.00 (1.65)		t(62) = -1.62, p = 0.11	
Years of education	18.35 (2.58)		17.00 (1.98)		t(62) = -2.34, p = 0.02*	
Length of residence in L2 environment (years)	9.66 (4.86)		-			
Length of reimmersion in L1 environment (days)	13.37 (8.18)		-			
Time delay between L1 reimmersion and recording (days)	3.06 (1.86)		-			
Self-assessed language experience (1-10)	L1	L2	L1	L2	L1	L2
Self-rated proficiency	9.82 (0.47)	7.86 (0.93)	9.7343 (0.70)	7.20 (1.29)	t(62) = -0.57, p = 0.57	t(62) = -2.33, p = 0.02*
Speaking	9.64 (0.79)	7.68 (1.25)	9.59 (1.07)	6.63 (1.52)	t(62) = -0.20, p = 0.84	t(62) = -3.02, p = <0.01*
Writing	9.77 (0.75)	7.42 (1.29)	9.72 (0.63)	6.75 (1.76)	t(62) = -0.28,	t(62) = -1.74,

					p = 0.78	p = 0.09
Listening	9.94 (0.25)	7.91 (0.86)	9.84 (0.51)	7.53 (1.29)	t(62) = -0.91, p = 0.37	t(62) = -1.37, p = 0.17
Reading	9.93 (0.25)	8.42 (0.98)	9.78 (0.79)	7.91 (1.35)	t(62) = -1.05, p = 0.30	t(62) = -1.74, p = 0.09
Percentage of daily use (%)	40.46 (15.94)	59.25 (15.53)	81.81 (16.19)	16.43 (11.13)	t(62) = 10.29, p = <0.01***	t(62) = -12.68, p = <0.01***
Age of L2 acquisition (years)	-	13.05 (3.72)	-	9.38 (3.82)	t(62) = -3.89, p = <0.01***	
Intensity of language switching (1 = no switching, 10 = always switching)	4.82 (2.55)		4.31 (2.26)		t(62) = -0.84, p = 0.40	
Objective L2 proficiency measures						
LexTALE (mean accuracy in %)	-	77.82 (13.24)	-	73.12 (10.97)	t(62) = -1.55, p = 0.13	
General English Test (mean accuracy in %)	-	89.88 (9.48)	-	84.50 (7.31)	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, (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 delay between the return from the L1 environment and the experimental recording. The second part of the table summarizes 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 the age of acquisition is 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 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 2.1), 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 when contacting friends and family in Poland, but never English or other languages. The language characteristics of both groups are summarized in Table 2.1.

4.3.2 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.

4.3.2.1 Materials

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 and created two versions with a different order of presentation for each subset. The subsets of pictures were balanced for name agreement (based on Wolna et al., 2022), lexical frequency (based on Mandera, Keuleers, Wodniecka, & Brysbaert, 2015), age of acquisition (Haman et al., 2015), and mean length in phonemes. 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.

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

4.3.2.2 Procedure

In the picture-naming task, pictures were displayed on a computer screen using DMDX (Forster & Forster, 2003). The pictures were presented in the center of the screen. Each trial was preceded by a black screen presented for 1000 ms, followed by a fixation cross in the screen's center for 1000 ms. A picture was then shown in the center of the screen until the participant responded or until the timeout was reached (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.

4.3.2.3 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 (0.05 – 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), down sampled to 256 Hz and baseline corrected. Ocular artefacts were removed with Independent Component Analysis (ICA, (Delorme et al., 2007; Jung et al., 2000) by calculating the ICA components based on 1 Hz high-pass filtered data and applying it to the 0.05 Hz high-pass filtered data set. Segments containing artefacts were cleaned manually. Afterwards, the data was exported to Matlab for further analysis using EEGlab (Delorme & Makeig, 2004) and ERPlab (Lopez-Calderon & Luck, 2014).

4.3.3 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: FC1, FC2, Fz, Cz, CP1, CP2 for the naming P2, and FC1, FC2, C3, C4, Cz, CP1, CP2 for the N300. The P2 electrodes were chosen based on Strijkers et al., 2013, as this study used the same setup with averaged mastoid references and investigated the effects of language use. Electrodes for N300 were based on Wodniecka et al., 2020. Additionally, due to temporal differences caused by the recording setup in each different location (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.

In order to be able to relate the electrophysiological results of the present study with the findings of the behavioral analysis already reported in Casado et al. (2023), we included in the analyses all the variables we controlled for in the previous study so as 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 participant’s variables the groups were not matched for: the participant's 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 (7.92%). We also excluded trials with inaccurate responses (2.55%). In total, 10.47 % of the data were excluded.

Table 2.2

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

	During L2 immersion/X-Context	After L1 reimmersion/Y-Context
Migrant group	88.3% (1525 trials)	81.7% (1411 trials)
Control group	86.4% (1493 trials)	82.0% (1417 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 centered 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.

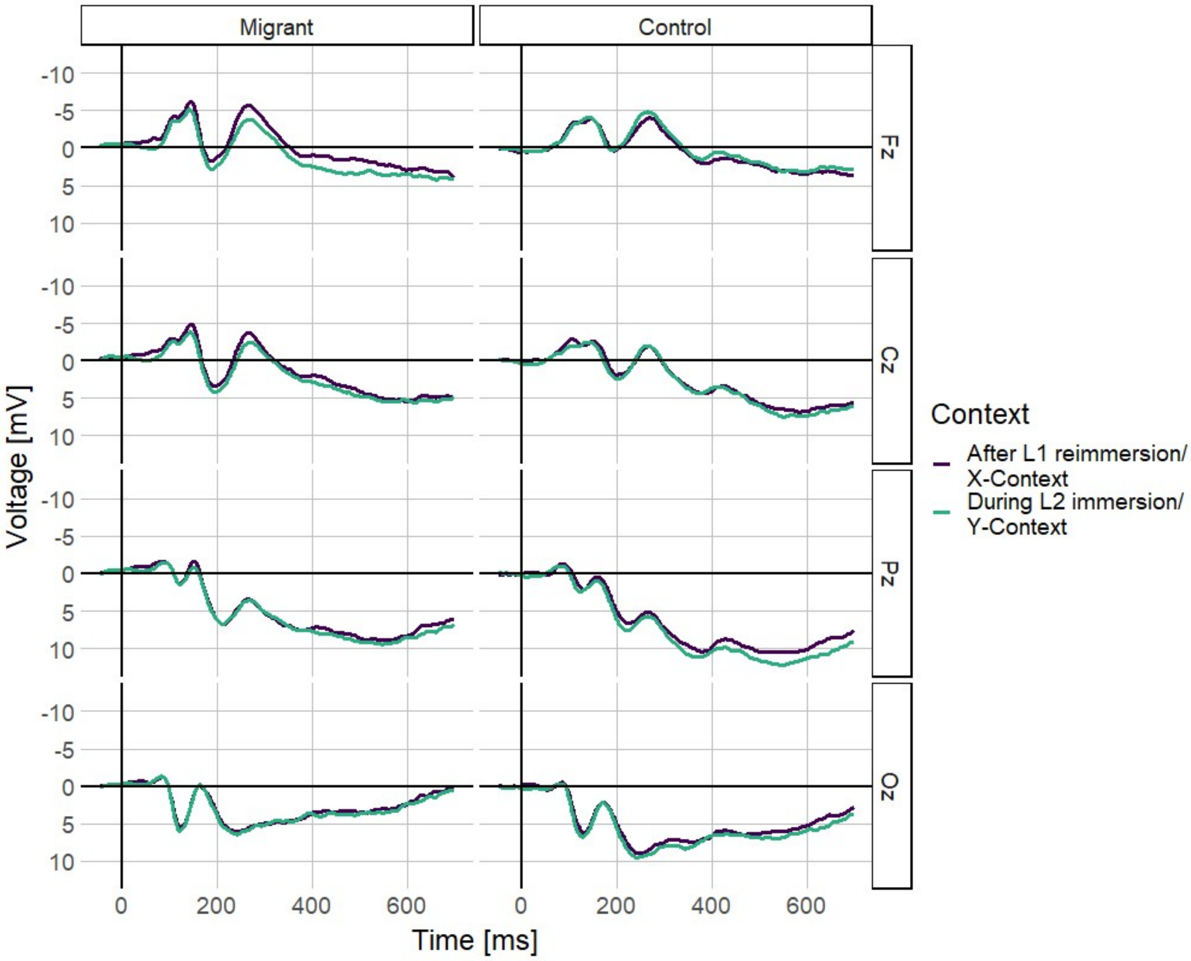
We fitted the maximal models first (Barr et al., 2013). If the model did not converge, we first removed correlations between random effects and in the next step, the random effects with the smallest unique variance, following the recommendation by Bates et al. (2018). Absolute t-values greater than two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

4.4 Results

The results are presented in Table 2.4 and Table 2.5. For the P2 component, the results revealed a significant interaction between Group and Context, demonstrating a modulation of the P2 component in the migrant group due to the language context: higher P2 amplitudes were observed during L2 immersion compared to after L1 reimmersion (see Figure 2.1). In contrast, no differences in the P2 amplitudes between contexts were found for the control group (see Table 2.3). Contrary to our predictions, we also did not observe the effect of frequency on mean amplitude in either the P2 or N300 components. Additionally, there was a significant effect of the participant's Age on the P2 amplitude with an increased amplitude for older participants. Overall, we found a more frontal distribution for the P2 within the L2 reimmersion effect than expected based on the previous studies (Costa et al., 2009; Strijkers et al., 2010; 2011; 2013).

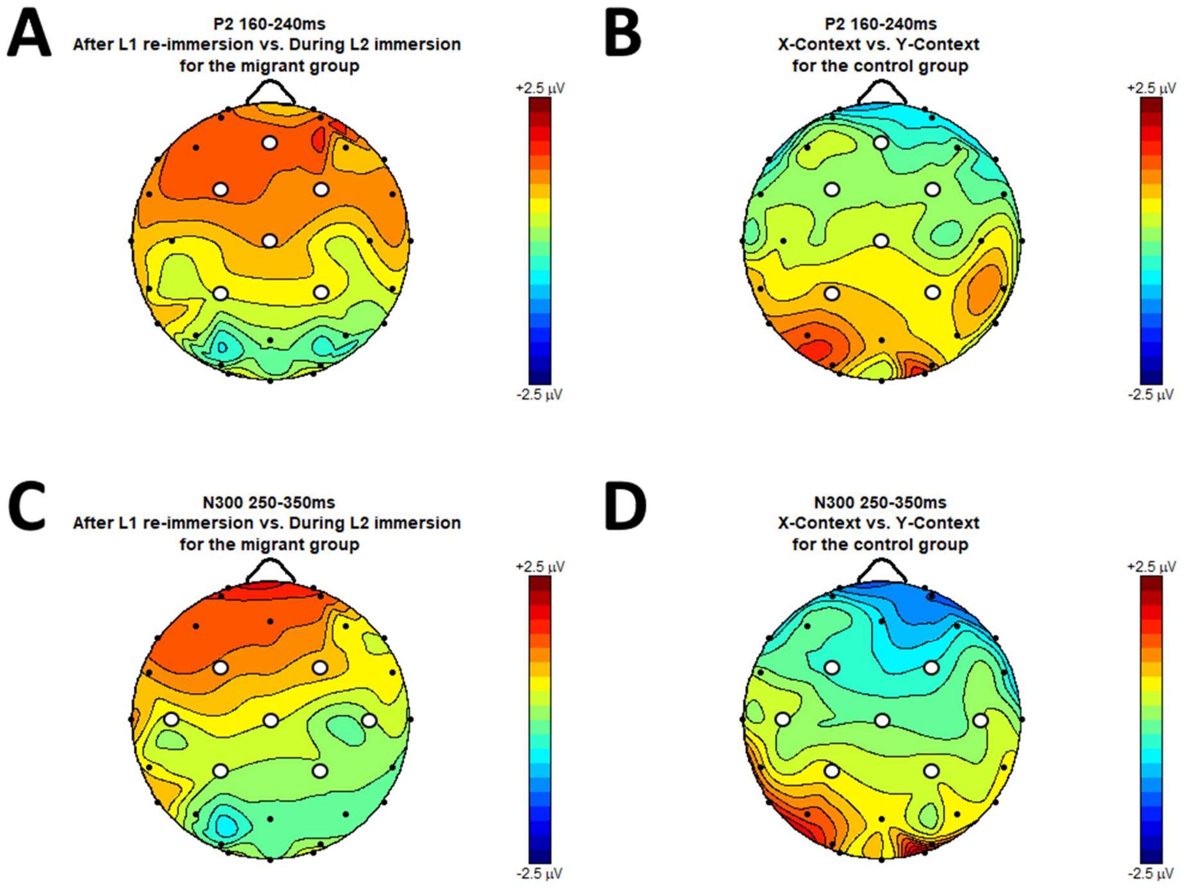
In the analysis of the N300 component, we found no main effects of Group or Context (see Figure 2.1 panel B) and no significant interactions.

Figure 2.1



Note. Stimulus-locked grand average ERP waveforms for naming P2 and N300 for the Migrant and Control groups. Waveforms depict mean voltages recorded from the midline electrodes Fz, Cz, Pz and Oz.

Figure 2.2



Note. Scalp topography maps for naming P2 **(A, B)** and N300 **(C, D)** and for the migrant group **(A, C)** and the control group **(B, D)**. Selected electrode clusters used for calculating mean amplitudes are marked in white.

Table 2.3

Summary of the electrophysiological data.

ERP component	Group	During L2 immersion / X-Context	After L1 reimmersion / Y-Context	Context comparison
P2	Migrant group	0.11 (0.07)	0.01 (0.07)	*z* = -2.10; *p* = 0.04*
	Control group	-0.09 (0.07)	-0.03 (0.07)	*z* = 1.42; *p* = 0.16
	Group comparison	*z* = -1.78; *p* = 0.08	*z* = -0.38; *p* = 0.71	
N300	Migrant group	-0.02 (0.08)	-0.10 (0.09)	*z* = -1.70; *p* = 0.09
	Control group	0.03 (0.09)	0.07 (0.09)	*z* = 0.76; *p* = 0.45
	Group comparison	*z* = 0.36; *p* = 0.72	*z* = 1.28; *p* = 0.20	

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.

Table 2.4

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	-0.07	0.07	-0.96	0.10	0.32
Group	0.12	0.10	1.15		
Context	0.01	0.03	0.47		0.18
Word-lexical frequency	0.00	0.01	-0.24	-	
Participant's Age	0.21	0.05	4.35***		-
Age of L2 acquisition	-0.05	0.05	-0.95		-
L2 proficiency	-0.06	0.04	-1.44		
log (Trial number)	0.02	0.02	1.19		
Group: Context	0.16	0.06	2.48*		-
Group: Word-lexical frequency	0.00	0.02	0.16		-
Control Group: Context: Word-lexical frequency	-0.01	0.03	-0.21		-
Migrant Group: Context: Word-lexical frequency	0.04	0.03	1.36		-

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

Table 2.5

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	-0.13	0.08	-1.53	0.22	0.39
Group	-0.11	0.13	-0.86		
Context	0.02	0.03	0.66	0.08	0.21
Word-lexical frequency	0.01	0.02	0.71	-	
Participant's Age	0.03	0.06	0.42		-
Age of L2 acquisition	-0.05	0.06	-0.78		-
L2 proficiency	0.02	0.05	0.33		
log (Trial number)	0.04	0.02	1.90		
Group: Context	0.12	0.07	1.75		-
Group: Word-lexical frequency	0.02	0.02	0.76		-
Control Group: Context: Word-lexical frequency	-0.01	0.03	-0.33		-
Migrant Group: Context: Word-lexical frequency	0.03	0.03	1.07		-

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

4.5 Discussion

In the present study, we explored neural signatures of L1 lexical access in language production of 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 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' 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 (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 and N300 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 and N300 component in migrants after a reimmersion in the L1 context compared to migrants during L2 immersion. Finally, if the effects of decreased access to L1 words are primarily visible in high-frequency words (as we argue in Casado et al., 2023), we should observe larger differences in frequencies for P2 and N300 amplitudes in the migrant group after reimmersion in L1 than during immersion in L2.

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 of overall differences between the groups in L1 naming latencies (see also Casado et al., 2023). We, therefore, conclude that despite the long-term immersion of the migrant bilinguals in the L2 environment, there are no clear traces of the loss of the native language access, either in overt behavior, or in the early electrophysiological markers of the lexical access. Still, the P2 component was sensitive to short-term changes in migrants' language environments, matching our second prediction. That is, higher amplitudes were evoked during L2 immersion vs. after L1 reimmersion, indicating that a short-term reimmersion in the L1 environment improved some aspects of the L1 lexical access in migrants, at least those detectable by the P2 component. However, we could not confirm our third prediction as there was no effect of the lexical frequency as both the P2 and N300 were insensitive towards any frequency manipulation. The N300 component was also not responding to any of our comparisons (migrants vs. control; migrants' L2 immersion vs. L1 reimmersion).

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.

4.5.1 Effects of long-term immersion

In line with our behavioral results reported in Casado 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 behavioral results which found that the lexical access in the native language of migrants living in L2 environments for 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 result in generalized difficulty in accessing native access words. Migrants are able to maintain L1 equally functional as in bilinguals who reside in the L1 environment on both the behavioral and neural levels, even after long-term L2 immersion.

The fact that access to L1 lexical information is not negatively affected by immersion in the L2 environment (as shown by the lack of group differences in naming latencies and in mean amplitudes) is consistent with the interface hypothesis originally proposed by Sorace and Filiaci (2006). This framework offers an explanation for the difficulties experienced by highly proficient balanced bilinguals immersed in an L2 environment when accessing L1. It posits that long-term L2 immersion only affects the sensitivity to high-level structures like the interface between syntax and pragmatics (e.g., during anaphora resolution, see Köpke & Genevska-Hanke, 2018; Chamorro et al., 2016; Chamorro & Sorace, 2016), rather than causing a permanent change in speakers L1 knowledge representations like vocabulary (Chamorro & Sorace, 2019; Sorace, 2011, 2016). Here we provide the first evidence from electrophysiological investigation suggesting that low-level structures (i.e., lexical level) indeed do not seem to be affected by L2 immersion. This adds 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).

4.5.2 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 the P2 amplitude was reduced. The P2 component was previously associated with the ease of lexical access with higher amplitudes under more difficult conditions (Costa et al., 2009; Strijkers et al., 2010, 2011, 2013). Following this interpretation, the decreased amplitude of P2 after L1 reimmersion seems to 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 P2 amplitudes compared to high-frequency words (Strijkers et al., 2010). Yet, in the present study, we did not find a modulation of the P2 component triggered by the word-lexical frequency either as a main effect (as expected based on Strijkers et al., 2013) or as interacting with the environment (as reported on the behavioral analysis presented in Casado et al., 2023). Therefore, it is unclear whether the observed P2 component should indeed be taken as an index of lexical access. Not only was our P2 not sensitive to word-lexical frequency, but also the overall distribution of our fronto-central component differs from the typical broad centro-parietal distribution reported in previous studies characterising the P2 as a lexical access index using lexical frequency (e.g., Costa et al., 2009; Strijkers et al., 2010; 2011) (see Figure 2.1 panel A, bottom). Indeed, Strijkers et al. (2013) too found differing distributions for the language effect (central, in L1 vs. L2 comparison) and lexical frequency (centro-parietal, for low vs. high frequency word comparison). Therefore, it is likely that the P2 component found in the present study does not reflect the difficulty of lexical L1 access, but a different underlying process.

One previous study using a picture-naming task reported a P2 component similar in terms of fronto-central distribution to the one reported here (Branzi et al., 2014) but offering a different interpretation than Strijkers et al. (2010) of what the P2 component indexes. Rather than seeing the P2 component as an index of lexical access difficulty, Branzi et al. (2014) suggested that P2 component could reflect the mechanisms of language regulation, 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 in L1 after using L2, higher 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 language control mechanisms applied during L1 lexical access to proactively manage the persisting activation of the previously used language —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 re-immersion in L1 influences the proactive control mechanisms applied during L1 lexical access. In our study, we found a modulation of the P2 component of the migrant group under the different language environments: P2 amplitudes were decreased after a reimmersion in the L1 environment compared to during immersion in the L2 environment. In our case, the higher P2 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 lower P2 amplitudes after reimmersion in the L1 environment could be reflecting 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 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 a 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 electrodes (Baus et al., 2020; Strijkers et al., 2010) can indeed reflect a lexical access difficulty that drives the word-frequency effect. On the other hand, the P2 component found over the fronto-central electrodes may reflect an 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 P2. First, in Strijkers et al. (2013), a larger 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 larger fronto-central P2 in response to naming in L1 after L2 compared to a baseline L1 naming. In this case, use of L2 prior to L1 can lead to a temporary increase in L2 activation level that results in increased interference between the two languages in the subsequent production in L1 (for the discussion of the influence of recent language use on bilingual language control see Blanco-Elorrieta & Caramazza, 2021; Casado et al., 2023). Finally, our results show increased fronto-central P2 in response to speech production during long-term immersion in L2 compared to a context of short-term re-immersion in L1. As previously explained, living in the L2 environment may result in an increase of L2 activation and consequently lead to higher interference between languages.

Regarding the N300, we decided to explore this component as it was previously identified as a marker of L1 lexical retrieval difficulties during picture recognition (Wodniecka et al., 2020). Still, we did not find modulations of the mean amplitude, neither during the between-group comparison nor when comparing migrants during L2 immersion and after L1 reimmersion. The lack of modulation in the between-group comparison aligns with the behavioral results (Casado et al., 2023), confirming that L1 retrieval abilities during picture recognition were not affected in migrants as a consequence of long-term immersion in the L2 environment. Moreover, the lack of modulation by the context in the migrant group supports the idea that the P2 component found in this study was not really capturing the ease of lexical access but could have been related to the appliance of proactive language control during L2 immersion vs. after L1 reimmersion. Still, additional studies are required to confirm this interpretation.

In summary, our results reveal that the way long-term migrants' access L1 words may be different depending on the language environment. Following Branzi et al. (2014) 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) studies, we propose that the P2 modulation observed in our study could reflect the engagement of proactive control in speech production. Under this explanation, involvement of proactive control applied during immersion in the L2 environment would be higher, as the level of L2 activation in this context is increased and thus controlling the interference between languages may be more effortful than in the L1 environment. It would get reduced following a short-term L1 reimmersion in which the L2 activation decreases, as L2 is used much less in the native language context. Altogether, our results could be taken as evidence for the hypothesis that bilingual cognitive systems adapt to language environments, adjusting strategies of language control as more or less proactive.

4.6 Conclusion

In this study, we show that migrants adapt their language 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 for the P2 component during L2 immersion and lower amplitudes after L1 reimmersion. We propose that this pattern of results demonstrates the use of different forms of language control depending on the language environment. 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.

5 Investigation 3: Free or forced language choice in language switching and its transfer to non-linguistic task switching - an ERP study

Jonas Walther[1], Kalinka Timmer[2], Patrycja Kałamała[1], Zofia Wodniecka[1]

1. Institute of Psychology, Jagiellonian University, Krakow, Poland
2. Faculty of Psychology, University of Warsaw, Poland

Acknowledgments

The research was possible thanks to a National Science Centre grant (2015/18/E/HS6/00428) awarded to Z.W. 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.

5.1 Abstract

Several previous studies showed transfer effects between language control and domain-general control. The present study investigated whether enhancement of cognitive control is driven by training in forced language switching (i.e., constrained by external cues) and would not occur after voluntary language switching. Participants completed a non-linguistic task switching paradigm before and after language switching training. Half of the participants were forced to switch languages while naming pictures, while the other half could choose the language freely on each trial. The behavioral results showed no significant change in switch costs from pre-test to post-test between the two groups. However, the forced language switching group revealed reduced mixing costs compared to the voluntary language switching group after training. A significant change in P3 mixing costs further supported this effect. The transfer of forced compared to voluntary language switching training to non-linguistic switching in mixing costs indicates that externally induced language switching requires sustained but not necessarily transient cognitive control.

5.2 Introduction

A remarkable feature of bi- and multilingual individuals is their ability to quickly and accurately switch back and forth between their first (L1) and second (L2) language. Switching between languages has been proposed to enhance non-linguistic cognitive control (i.e., the set of cognitive processes responsible for goal-directed behavior; for an overview, see Friedman, 2016; Kroll et al., 2014). However, the idea of shared mechanisms between language switching and cognitive control has come under criticism recently, with inconsistent evidence for and against it (Lehtonen et al., 2018). Recently, Blanco-Elorrieta and Pylkkänen (2018) have hypothesized that language switching can only enhance cognitive control if it occurs under certain communicational conditions. Only if language switching is forced by external demands, such as a monolingual interlocutor, does it involve and thereby enhance cognitive control. Following similar ideas, we designed a training study that compared the impact of forced versus voluntary language switching on non-linguistic task switching efficiency. Polish-English bilinguals completed a picture-naming training involving two types of language switching. Before and after the training, they performed a non-linguistic switching task to measure the impact of language switching training on cognitive control mechanisms. Based on the Adaptive Control Hypothesis (ACH hereafter) by Green and Abutalebi (2013), we expected a greater reduction in non-linguistic switch costs after forced language switching training than after voluntary language switching training. To supplement the behavioral data, we also recorded brain activity using electroencephalography (EEG), specifically the P3 component to measure mental effort during task switching.

5.2.1 The Adaptive Control Hypothesis

The ACH suggests that the effects of bilingualism depend on the specific language usage that a speaker engages in their daily life. Depending on the individual experiences, the associated control processes adapt to the demands imposed on them by the language context. Specifically, the hypothesis distinguishes between three primary conversational contexts: a single-language context, where a single language is limited to a single context, a dual-language context, in which two languages are presented but separated for different speakers, and finally, the dense code-switching context, wherein two languages are mixed within speech and single utterances.

The ACH makes predictions for eight different cognitive control abilities used differently in the three contexts. The most interesting discrepancies exist between the dual-language and code-switching contexts. Because in the dual-language context, speakers have to be ready to switch their language immediately when facing a new speaker, among other abilities, cue detection, response inhibition, task engagement, and disengagement are required. Speakers must be able to recognize and process the arrival of new dialogue partners to know when they have to switch to a different language, activating their cue-detection abilities. When a new conversation has been established, they have to stop themselves from using the incorrect language and engage in response inhibition to suppress any spurious activations of non-intended language. Finally, task engagement or disengagement is required to complete the entire process of activating a new required language system while shutting down the old one.

In the dense code-switching context, carefully selecting a given language is less important as all speech partners can understand multiple languages and mix their languages freely within conversations and utterances. Therefore, a speaker in the dense code-switching context should have a reduced need for cue detection, inhibitory control, or task engagement and disengagement. Instead, they can optimize their speech patterns by flexibly adapting to the present situation using previously primed phrases or selecting words based on their dynamic accessibility.

5.2.2 Testing the ACH

Because the language contexts described above vary greatly in the required executive control needs, speakers who are active in different conversational contexts should also, according to the ACH, demonstrate varying efficiency in cognitive non-linguistic tasks that share underlying cognitive mechanisms. For example, bilinguals recently having been in a dual-language context should outperform speakers with recent exposure to the code-switching context in tasks involving response inhibition like the go/no-go task due to their extensive use of response inhibition in their language use in the dual-language context.

There are two possible methodological approaches to test the different usage of cognitive control between conversational contexts. The first is to compare the performance of bilingual populations with differences in their daily language use and language switching behavior in non-linguistic cognitive tasks. With such a design, Beatty-Martínez et al. (2020) compared Spanish-English bilinguals from an L1-dominant environment, an L2-dominant environment, and a code-switching environment. The authors found that the three groups differed in their use of cognitive control, as measured by the AX-CPT task. While bilinguals in their native language environment relied more on adaptive strategies, bilinguals from an L2-dominant environment used more proactive control, while code-switching bilinguals used a mix of these strategies. Similar between-group comparisons found that bilinguals from the dual-language context, compared to single-language context bilinguals, showed reduced task switch costs (Hartanto & Yang, 2016) and more efficient conflict resolution (Ooi et al., 2018).

The second approach involves the experimental simulation of language contexts and assessing their impact on non-linguistic cognitive tasks. Two typical cognitive tasks in these types of experiments are the linguistic switching task using picture naming and the non-linguistic task switching task. The picture naming task involves presenting a series of pictures to participants accompanied by a cue that indicates the appropriate language in which a picture should be named (i.e., language switching paradigm). By using multiple cues for different languages, participants are asked to switch between languages in a controlled setting resembling the dual-language context. Typically for this task, longer naming latencies are observed when switching between languages compared to repeating the same language (switch cost). In addition, when switching from the weaker language (L2) to the more dominant one (L1), naming latencies are also increased (switching asymmetry, Meuter & Allport, 1999). Finally, participants often demonstrate longer naming latencies in blocks of language switching compared to pure language blocks in which only one language is used (mixing costs, Christoffels et al., 2007). While switch costs are hypothesized to reflect short-term or transient language control mechanisms that operate on a trial-to-trial basis, mixing costs are thought to manifest more global control processes that are applied to the entire or large parts of the language system, such as conflict monitoring (for a review see: Declerck & Philipp, 2015). Non-linguistic task switching follows a similar design as the linguistic switching task. Participants categorize presented shapes according to color or shape rules during this paradigm. The sorting rule to be followed by the participant is determined by a given cue, for instance, if they should sort stimuli by color or shape. Like language switching, participants also demonstrate switch and mixing costs with longer reaction times in switch trials than repetition trials and longer reaction times in mixed blocks than single-task blocks. While switch costs are assumed to reflect the after-effects of task-set reconfiguration, mixing costs are thought to reflect global cognitive processes and differences in working memory load (for a review, see: Kiesel et al., 2010). Given these well-known effects, multiple studies have used language switching (Zhang et al., 2015), non-linguistic task switching (Hartanto & Yang, 2016; López Zunini et al., 2019; Paap et al., 2017), or both (Calabria et al., 2012; Liu et al., 2019; Prior & Gollan, 2013; Timmer et al., 2019) to examine the relationship between language control and domain-general control. Several of these studies found transfers between language switching training and non-linguistic domains. For instance, participants showed better conflict monitoring by reducing mixing costs in non-linguistic switching after naming Arabic numerals in either Mandarin or English (Liu et al., 2019). Training in Chinese-English picture naming has also been shown to improve performance in the AX-CPT task, suggesting improved goal maintenance (Zhang et al., 2015). Similarly, training in cued non-linguistic task switching can transfer to a cued language switching task using digit-naming in the form of reduced mixing costs (Prior & Gollan, 2013). Crucially, the transfer effects differ for the type of language training. When comparing naming in only L2 against switching in both languages, pure block naming in L2 can improve inhibitory control (Liu et al., 2019) and language switching training has been shown to lead to greater efficiency in non-linguistic switching (Timmer et al., 2019) and conflict monitoring (Liu et al., 2019). Timmer et al. (2019) examined two groups of bilinguals who completed a non-linguistic switching task in a pre- and post-training design. One group of bilinguals received language switching training during the training session with cues determining the language, and another group received blocked naming training. The results revealed that training in language switching, but not blocked naming, decreased switch costs in the post-training session, demonstrating the close relationship between language specific processes and domain-general cognitive control.

Still, there remain some open questions. Many of the studies mentioned above contrasted language switching training to either no training (Prior & Gollan, 2013; Zhang et al., 2015) or monolingual picture naming (Liu et al., 2019; Timmer et al., 2019). Therefore, it is unclear where the observed transfers from language switching to cognitive control originate. Are they the result of simply using multiple languages within a mixed language context, or are they caused, as the ACH suggests, by some unique environmental conditions of the language context? The question is if an increased transfer from linguistic to non-linguistic domains of specific cognitive control abilities arises because of the presence of two languages or because higher demands were placed on the cognitive control system through unique properties of the language context.

5.2.3 Language competition and cooperation

Following the ACH, Blanco-Elorrieta, and Pylkkänen (2018) made more specific predictions for conditions under which language switching might impact non-linguistic cognitive control abilities. They argued that experimental paradigms must be carefully selected to create a realistic and ecologically viable simulation of the bilingual experience. They observed that language switching is not as effortful as language switching within artificial cognitive tasks in many real-life cases. The researchers proposed that only forced language switching based on external constraints is effortful and requires additional executive control. In contrast, voluntary language switching, in which participants can choose their language freely, shows significantly reduced costs, processing speed, and cognitive resources (Blanco-Elorrieta & Pylkkänen, 2017). While participants need to follow a given language instruction during forced language switching, during voluntary language switching, bilinguals can adjust their language system based on internal demands, like lexical accessibility and semantics, requiring fewer cognitive resources. Because voluntary language switching requires less cognitive control, these cognitive mechanisms are less activated, and there is no performance transfer from the linguistic to the domain-general cognitive control.

Following the ACH, it seems likely that language contexts place different levels of demand on the cognitive system, impacting the executive control system to varying strengths. Only contexts with relatively high cognitive demand can train executive control abilities and lead to improvements in general cognitive control. In their recent review, Beatty-Martínez and Titone (2021) followed a similar line of thinking. They suggested that two languages can be used cooperatively, in contrast to contexts where the two languages are either highly compartmentalized and compete with each other or where the two languages are mixed only in a highly regulated way. In that sense, language cooperation refers to environments that allow the beneficial use of multiple languages within the same context.

Evidence from so-called voluntary language switching paradigms supports the idea that language switching does not have to be adversarial and reduce the speed of using one or both languages. This task is in many ways identical to the language switching task using picture naming, only that there is no cue present to determine the target language used by the participant. Instead, participants are asked at the beginning of the task to use both languages during the task. Participants are, therefore, able to choose languages of their own volition. Some initial studies examining voluntary language switching have already found performance advantages in overall speed and switch costs during language production. Generally, participants were faster as they could choose words freely based on lexical access and word frequency to optimize their speech production (de Bruin et al., 2018; Gollan & Ferreira, 2009; Gross & Kaushanskaya, 2015). By using each language depending on the availability of a given word in L1 or L2, bilinguals were able to reduce switch costs and decrease naming latencies overall with, in some cases, even greater speed than monolingual speech (de Bruin et al., 2018; Gollan et al., 2014; Jevtović et al., 2020; Kleinman & Gollan, 2016). Furthermore, previous studies demonstrated that voluntary language switching leads to reduced activation of brain areas related to language control (Blanco-Elorrieta & Pylkkänen, 2017), suggesting that cooperative language switching requires less cognitive effort to control.

It remains an open question what the exact processes are underlying forced versus voluntary language switching. The ACH predicts that only language switching controlled by external constraints, like language cues, requires extensive control processes (task engagement and disengagement), while voluntary language switching primarily uses opportunistic planning. In a similar line of thought, Blanco-Elorrieta and Pylkkänen (2018) suggest that only external demands, such as language cues forcing a language switch, would lead to a transfer from language switching to non-linguistic task switching performance. Our study aims to test whether external constraints on language choice (i.e., cues incorporated in the task) involve additional domain-general cognitive control. We compared forced and voluntary language switching and its impact on non-linguistic switching in behavioral and electrophysiological data.

5.2.4 Electrophysiological measures of task switching and mixing

A popular measure to assess non-linguistic task switching performance is the target-locked central-parietal P3 component. It is a positive event-related potential around 300ms after stimulus presentation. The typical pattern in the target-locked P3 is a more negative amplitude on switch compared to repeat trials (Karayanidis et al., 2003, 2011; Kieffaber & Hetrick, 2005; Lorist et al., 2000; Nicholson et al., 2005; Poulsen et al., 2005). During task mixing, the P3 also shows a more negative amplitude in mixed block trials compared to pure block trials (Bae & Masaki, 2019; P. D. Gajewski & Falkenstein, 2011; Goffaux et al., 2006; Jost et al., 2008).

Generally, a more negative amplitude for the P3 component is assumed to reflect fewer available resources in working memory to process the current target (Kok, 2001; Polich, 2007). For switch costs, a more negative P3 amplitude during switch trials was interpreted as higher task demands and an increased working-memory load, resulting in a greater task difficulty for switch than repeat trials (Karayanidis et al., 2003, 2011; Lorist et al., 2000). In a similar way of thinking for mixing costs, more negative P3 amplitudes during repeat trials from the mixed block compared to trials from the pure block reflect the need for greater working memory capacity to maintain multiple task rules in mind at the same time (Goffaux et al., 2006; Kok, 2001). The amplitude of the P3 can, therefore, be seen as an indicator of mental effort: a more positive P3 amplitude is interpreted as a high amount of resources available to a single task, while a more negative P3 amplitude indicates a lower amount of resources available, either a result of competition between multiple tasks or an increase in single task difficulty (Lorist et al., 2000; Wickens et al., 1983).

However, this interpretation of P3 is not always as straightforward. Some studies (Barcelo et al., 2006; Timmer et al., 2017) found a more positive P3 amplitude for switch trials than for repeat trials. A possible explanation for these results might be that P3 reflects a multitude of different mental processes that contribute differently to the overall cognitive demands indexed by P3 (Kok, 2001). Depending on the specific task design, mental processes such as attention and working memory might independently increase switch and mixing costs. While the P3 amplitude might increase due to the availability of additional mental resources during pure block trials in task mixing, a greater task difficulty of switch trials during task switching might decrease the P3 amplitude. For instance, the target-P3 amplitude might increase in pure block trials compared to repeat trials from the mixed block as a result of the greater availability of mental resources. Within the same experiment, other factors such as age or exposure to mental exercises could decrease it (Gajewski & Falkenstein, 2012). Therefore, the totality of the interplay between these processes can determine the final P3 amplitude patterns. Which direction the P3 amplitude changes take could consequently depend on the involved processes, which are determined by the specific task design (Kok, 2001).

Table 1

Overview of non-linguistic task switching studies reporting a target-locked P3 component

Study	Participants (N)	Time window (ms)	Distribution	P3 amplitude for switch vs. repeat trials	P3 amplitude for pure vs. repeat trials
Bae and Masaki, 2019	30	280 – 380	Parietal	Repeat more positive	Pure more positive
Barcelo et al., 2006	16	330 – 350	Frontal-central	Switch more positive	-
Gajewski et al., 2010	91	300 – 600	Centroparietal	Repeat more positive	Pure more positive
Gajewski and Falkenstein, 2011	17	300 – 600	Centroparietal	Repeat more positive	-
P. Gajewski and Falkenstein, 2012	141	300 - 600	Centroparietal	No effect	Pure more positive
Gehring et al., 2003	14	300 – 500	Central	Repeat more positive	-
Goffaux et al., 2006	20	300 – 600	Posterior	Repeat more positive	Pure more positive
Hsieh and Liu, 2009	12	300 – 600	Centroparietal	Repeat more positive	-
Jost et al., 2008	16	300 – 800	Parietal	-	Pure more positive
Kamijo and Takeda, 2010	40	250 – 600	Parietal	No effect	Pure more positive
Karayanidis et al., 2011	95	400 – 600	Parietal	Repeat more positive	Pure more positive
Kieffaber and Hetrick, 2005	39	-	Anterior	Repeat more positive	-

López Zunini et al., 2019	79	300 – 600	Centroparietal	Repeat more positive	Pure more positive
Lorist et al., 2000	16	200 – 700	Anterior	Repeat more positive	-
Nicholson et al., 2005	24	400 – 600	Parietal	Repeat more positive	-
Periáñez and Barceló, 2009	41	310 – 730	Parietal	Switch more positive	-
Petruo et al., 2019	52	320 – 370	Parietal	Repeat more positive	-
Poulsen et al., 2005	21	482 – 570	Parietal	No effect	-
Travers and West, 2008	20	200 – 400	Parietal	No effect	Mix more positive
	20	200 – 400	Parietal	No effect	Mix more positive
Richardson et al., 2022	22	300 – 500	Parietal	Repeat more positive	Pure more positive
Tsai et al., 2016	60	250 – 600	Parietal	Repeat more positive	Pure more positive
Timmer et al., 2017	43	325 – 400 (switch) 300 – 375 (mix)	Central to occipital	Switch more positive	Pure more positive
Whitson et al., 2014	95	300 – 700	Centroparietal	Repeat more positive	Pure more positive
Wolff et al., 2018	26	320 – 380	Parietal	Switch more positive (random trial order) Repeat more positive (fixed trial order)	-
Wong et al., 2018	201	200 – 600	Centroparietal	Repeat more positive	Mix more positive

Note. Summary of studies examining target-locked P3 in non-linguistic task switching paradigms**.**

5.2.5 Current study

The present study investigated how forced and voluntary language switching impact non-linguistic task switching performance. Following previous claims in the literature, we hypothesized that forced language switching should set greater demands on the cognitive control system than voluntary language switching and, therefore, better train aspects of cognitive control related to task-set reconfiguration. If that is the case, we should be able to observe the differential consequences of two types of training on subsequent non-linguistic task switching performance. As the forced language switching group is encouraged to maintain the correct language use due to external constraints, the group should have a greater need for transient cognitive control. The voluntary switching group would be able to switch based on internal preferences, such as lexical access, and could therefore rely to a lesser degree on cognitive control. This would lead to a greater transfer in shared mechanisms between language switching and non-linguistic task switching in the forced language switching group compared to the voluntary language switching group. More specifically, we hypothesized that, by transferring the shared underlying processes between the linguistic and non-linguistic domains, the forced language switching training should lead to greater transfer reflected by increased efficiency during subsequent task switching in the form of reduced switch costs than the voluntary language switching training. The idea is that switch costs are assumed to represent these transient cognitive processes dealing with the interference from two distinct task-sets (Kiesel et al., 2010). Mixing costs, on the other hand, are thought to represent differences in global cognitive control and the availability of mental resources for maintaining multiple task sets (Braver et al., 2003; Prior & Macwhinney, 2010). Both groups will use L1 and L2 to similar degrees during language training, so there should be no differential transfer to the non-linguistic task switching afterward. Using a similar experimental design to Timmer et al. (2019), we investigated two matched groups of bilinguals who completed two different types of language switching training. One group used external cues to indicate language switches, while the other was able to switch between languages voluntarily. Comparing the group's performance in a non-linguistic task switching task before and after training allowed for the measurement of the different impacts of dual-language context and dense code-switching context on domain-general cognitive abilities active during task switching.

In preparation for the main experiment, we conducted two pilot studies: Pilot 1 and 2. In Pilot 1, we established values for name agreement and language preference for our stimuli in our target population. For Pilot 2, we selected stimuli based on the results of Pilot 1. We designed a voluntary language switching task to determine the average switching behavior of our participants.

For domain-general cognitive control, we focused on task switching and task mixing during the non-linguistic task switching paradigm. For switch costs, the forced language switching group requires greater cognitive control and should benefit from greater efficiency transfer between the tasks. Mixing costs, on the other hand, are determined by the difficulty of managing multiple activated languages, while language switching requires the engagement and disengagement process on a trial-to-trial basis. Because the two types of language training should involve similar levels of activation of Polish and English and, therefore, similar levels of global language control, we expect no differential transfer of mixing efficiency between the linguistic and non-linguistic tasks for the two participant groups. The measurement of both switch and mixing costs in the non-linguistic task switching task allowed us to dissociate the effects of externally controlled language switching and the overall exposure to two languages. To better assess the availability of cognitive resources during non-linguistic task switching, we also supplemented the behavioral measurement by recording the target-locked P3 component to gain additional information about mental processing during non-linguistic task switching.

Suppose external demands during language switching in the form of language cues significantly impact the use of cognitive mechanisms. In that case, forced and voluntary language switching should differ in the extent to which they involve domain-general cognitive control. This difference should be reflected in transferring the general switching efficiency to a non-linguistic task switching paradigm. Following these assumptions, we proposed the following predictions for the main experiment:

Prediction 1: After the forced language switching, participants will show smaller switch costs in reaction times, accompanied by smaller switch costs in the P3 mean amplitude, than after the voluntary language switching.

This prediction is based on the idea that using language cues during the forced language switching restricts language choice within each trial and engages additional cognitive control abilities to maintain activation of the required language and prevent incorrect answers. The use of cognitive control will therefore be prominent in the forced language switching group and lead to a greater training effect compared to the voluntary language switching group.

Prediction 2: Both forced and voluntary language switching should lead to comparable mixing costs observed in the subsequent task switching: no group differences were expected for mixing cost indices (repeat vs. pure trials) either in reaction times or P3 mean amplitude.

We formulated this prediction based on the idea that while language cues restrict language choice for a particular trial, they do not change the overall presence of two languages in a given testing block, and therefore the global activation of both languages should be the same under the forced and voluntary language switching block. In other words, both types of language switching require similar global language control to manage language mixing and should show similar changes in mixing costs.

5.3 Methodology

Pilot 1. Language preference pilot study

Pilot 1 aimed at identifying pictures with high language preferences for the targeted languages (i.e., Polish and English). To this end, participants indicated whether Polish or English came first to their minds when seeing a set of pictures. The data served to calculate a language preference index. For the same pictures, participants completed a picture-naming task which served to calculate the name agreement for each picture. Additionally, each participant completed an L2 proficiency test and a background questionnaire.

Pilot 2. Voluntary-switching pilot study

Pilot 2 aimed to test whether the pictures selected based on language preference index and name agreement (Pilot 1) could elicit voluntary language switching in Polish-English bilinguals. The participants completed a version of the picture-naming task in which they were instructed to switch between languages on a voluntary basis. We assessed whether language choice adhered to the language preference index and whether classic language switching asymmetry occurred based on the collected data.

Main experiment. Cognitive-demand study

The main experiment aimed to test the conditions under which language switching enhances domain-general cognitive processes. The participants completed a non-linguistic task switching task before and after language training. For one half of the participants, the language training took the form of voluntary language switching, and for the other half, forced language switching. We recorded both behavioral and electrophysiological (EEG) measurements. The comparison of neurocognitive performance between the two types of language training allowed the investigation of how different types of language switching affect the mechanisms of cognitive control.

5.3.1 Methodology Pilot 1

5.3.1.1 Participants

Only right-handed Polish native speakers in the age range of 20 – 35 years who lived in Poland and declared a relatively good proficiency in English were invited to the study. The total number of participants was 60, all recruited using the online platform Prolific (www.prolific.com; for the questions used for the participants' selection, see Appendix 8.3.1). All participants were healthy young adults (20 females; mean age of 23.48 years, *SD* = 3.98) with normal or corrected-to-normal vision. The participants considered themselves as having average socioeconomic status[5].

All participants were late language unbalanced bilinguals. They acquired Polish as their native language (L1) from birth and, on average, started learning English (L2) around the age of 6.73 years (*SD* = 1.99). The participants' language proficiency and use were evaluated using a language background questionnaire based on Marian and colleagues (2007) and Li and colleagues (2014). They rated their L1 proficiency (*M* = 8.74, *SD* = 1.11) significantly higher than their L2 proficiency (*M* = 7.51, *SD* = 1.06; *p* < .001). Descriptive statistics related to language proficiency can be found in Appendix 8.3.2. Their proficiency in L2 was additionally measured with a lexical test for advanced English learners (English LexTALE hereafter; Lemhöfer & Broersma, 2012). On average, participants scored relatively highly in this task (*M* = 71.6%; *SD* = 11.6%). The data of all participants were included in the reported analysis.

5.3.1.2 General procedure

Pilot 1 was conducted online via the Gorilla platform (www.gorilla.sc). It consisted of four tasks and one questionnaire, always administered in the following order: (1) a language preference

---

[5] The participants' socioeconomic status was assessed using an *8*-point scale based on Adler et al. (2000) from *1* for "I see myself at the lowest level of society" to *8* for "I see myself at the highest level of society". Their average score was 4.68 (*SD* = 1.26) which referred to average socioeconomic status.

task, (2) a vocal picture-naming task, (3) a written picture-naming task, (4) a language background questionnaire, and (5) the LexTALE. Participants performed the tasks on their personal computers with all instructions given in Polish on the screen. As the vocal picture-naming task required participants to name pictures aloud, we first verified the audio quality of their data with a short practice session (five trials to be named in Polish). Only participants with good quality vocal recordings were invited to participate in the pilot. The completion of all tasks took approximately 50 minutes.

5.3.1.3 Stimuli

The picture set for the language preference, vocal-, and written picture-naming tasks was the same. It consisted of 96 colored pictures from the database for Cross-Linguistic Lexical tasks (Haman et al., 2017) and 22 black-and-white pictures from a study by Blanco-Elorrieta & Pylkkänen (2017), resulting in 118 pictures in total.

All pictures depicted non-cognates and non-compound nouns with a name agreement score above 80% in both Polish and English (Szekely et al., 2004). We selected pictures with the highest word frequency differences between Polish and British English under the assumption that the more often encountered a word in a language, the greater the preference for the use of that language. The word frequency was established based on the SUBTLEX databases for Polish (Mandera et al., 2015) and British English (Heuven et al., 2014). The selected databases provided two advantages over the other available databases. First, word frequency derived from subtitles from TV shows and movies instead of literature is considered to give a better estimate of natural word use on a daily basis. Second, subtitles have been shown to provide a better-standardized measure of word frequency, i.e., the Zipf-scale (i.e., the decadic logarithm of the frequency of a given word per billion words in the corpus; for details, see Heuven et al., 2014).

5.3.1.4 Procedure and design

5.3.1.4.1 Language preference task

Participants were asked to assess their language preference for every picture from the set of 118 pictures (see section 5.3.1.3.). They indicated in which language (i.e., Polish or English) they preferred to name each displayed picture. In every trial, participants first saw a fixation cross for 250 ms, followed by the presentation of a picture and a slider, one below the other. The slider was labeled with “Polski” (“Polish”) on the left and “Angielski” (“English”) on the right side. The slider’s initial placement was in the center position, and participants moved the slider to the left for Polish and to the right for English preference. The slider allowed answering on a continuous scale, and no numerical values were presented. The farther from the center the slider was, the greater the preference for using a given language. There was no time limit for responding, and the participants advanced to the next trial on their own time with a button press. The answers of the participants were recorded as a numerical value between -100 and 100, with -100 for Polish preference and 100 for English preference (language preference index hereafter).

5.3.1.4.2 Vocal picture-naming task

Participants were asked to name vocally every picture from the set of 118 pictures (see section 5.3.1.3.). The pictures were presented one by one. Each picture was shown once for Polish and once for English, adding up to 236 trials in total. Each trial started with a 250 ms fixation cross, followed by the presentation of a language cue (i.e., Polish flag for Polish naming or British flag for English naming) and the picture together for the fixed interval of 3000 ms. The picture was displayed on the screen for the given time regardless of when a response was given. The order of the pictures and the two language cues were fully randomized. The task was divided into four mixed blocks (59 trials each), and the participants could take short breaks between each.

5.3.1.4.3 Written picture-naming task

Participants were asked to write down the names of all 118 pictures in Polish and English in the corresponding text fields. Pictures were displayed one by one in random order. Two text fields were presented below the picture. Participants could advance to the subsequent trial at a self-paced speed without a time limit. The purpose of this task was to verify the answers given in the vocal naming task in case of bad audio quality. Since no problems with audio quality were encountered, only the vocal picture-naming data was analyzed.

5.3.1.5 Analysis

The analyses aimed at calculating the name agreement and language preference index for all 118 pictures. For the vocal picture-naming task, only responses that were understandable and consisted of the correct name or synonym were included in the analysis. The name agreement scores for Polish and English were calculated by dividing the number of the most common picture name by the sum of the correct responses for each picture separately. For assessing language preference in Polish and English, the language preference index for each picture in each language was averaged across all participants.

5.3.1.6 Results

The mean name agreement scores were 93.95% for Polish names (*SD* = 10.45%; range: 49.12% - 100%) and 86.26% for English names (*SD* = 26.49%; range: 35.29% - 100%).

The mean language preference index across all pictures was -20.13 (*SD* = 28.20; participant range: -88.04 – 47.41), favoring on average Polish over English.

5.3.2 Methodology Pilot 2

5.3.2.1 Participants

The participants' recruitment procedure and criteria were the same as in Pilot 1 (right-handed, Polish native speakers with a high English proficiency between the ages 20 to 35 years, normal or corrected-to-normal vision), with the additional criterion of a score of at least 60% on the English LexTALE (Lemhöfer & Broersma, 2012). Sixty Polish-English bilinguals were recruited for Pilot 2. Out of the total sample, nine participants were excluded from the reported analysis due to a low number of switching trials (less than 5%). Therefore, the analyses were conducted on a final set of 51 participants (13 female; mean age of 23.34 years, *SD* = 3.32). The participants considered themselves as having average socioeconomic status (SES rating: *M* = 4.78, *SD* = 1.33; for rating scale, see Footnote 1).

All participants were late language unbalanced bilinguals. They acquired Polish as their native language (L1) from birth and learned English (L2) at a mean age of 6.48 (*SD* = 3.56). Similar to Pilot 1, they rated their L1 proficiency (*M* = 8.93, *SD* = 1.18) as significantly higher than their L2 proficiency (*M* = 7.61, *SD* = 1.35; *p* < .01). The proficiency in L2, measured with LexTALE, was relatively high (*M* = 71.63%; *SD* = 8.40%). Detailed descriptive statistics related to language proficiency can be found in Appendix 8.3.2.

5.3.2.2 General procedure

The study was conducted online via the Gorilla platform (www.gorilla.sc). Similarly, to Pilot 1, we first verified the audio quality of the data for the participants with a five-trial picture-naming task in Polish. Only participants with good quality vocal recordings were invited to participate in the pilot. During the pilot, participants completed a voluntary language switching task followed by a sociodemographic and language background questionnaire. All instructions were given in Polish on the computer screen. The entire pilot took approximately 35 minutes.

5.3.2.3 Stimuli selection

A group of 16 pictures was selected based on the language preference data collected in Pilot 1. Only pictures with a name agreement score above 75% in Polish and English were included. We only chose stimuli with different name onset letters and similar word lengths between translation equivalents (stimuli with a difference of four letters or more in their names were excluded). Based on the language preference index from Pilot 1 (for a detailed description, see section 5.3.1.4.1.), we created three stimuli sets: high Polish preference, high English preference, and ambiguous language preference. For each stimulus with a given preference for Polish, we have matched the stimuli with the equivalent preference for English. For example, if a stimulus had a relatively high preference for Polish (corresponding to a value of -20 on the slider), we matched the stimuli with a relatively high preference for English (corresponding to a value of 20 on the slider). Eight additional stimuli with ambiguous preference (corresponding to 0 on the slider) were also selected. Out of 16, four stimuli had high Polish preference (M = -19.63, SD = 6.86), four had high English preference (M = 22.18, SD = 5.74), and eight had ambiguous language preference (M = -1.15, SD = 5.63). The stimuli were also controlled for word frequency in Polish, American, and British English and semantic relatedness in English between the stimuli sets (Table 3.1; ps > .05).

Table 3.1

Linguistic properties of the three stimuli sets.

Statistic	Stimuli sets		
	Polish preference	Ambiguous preference	English preference
Mean language preference index[a]	-19.63 (6.86)	-1.15 (5.63)	22.18 (5.74)
Word frequency Polish[b]	3.96 (0.74)	4.24 (0.76)	3.97 (0.12)
Word frequency British English[b]	4.37 (0.36)	4.76 (0.56)	4.43 (0.23)
Word frequency American English[b]	4.47 (0.23)	4.73 (0.56)	4.50 (0.14)
Word Length Polish[c]	6.00 (0.82)	5.13 (1.25)	6.00 (2.94)
Word Length English[c]	4.25 (0.95)	5.00 (0.75)	4.75 (1.25)
Semantic relatedness English[d]	0.19 (0.04)	0.19 (0.03)	0.22 (0.03)

Note.

[a] Mean language preference index is based on the results of Pilot 1, with -100 indicating a preference for Polish and 100 for English (for details, see section 5.3.1.4.1).

[b] Word frequencies in Polish, British and American English are based on the Zipf scale (see Heuven et al., 2014).

[c] Word length in Polish and English denotes the number of characters in each stimuli name.

[d] Semantic relatedness was calculated using word2vec (Mikolov et al., 2013) trained on the Google News corpus (https://code.google.com/archive/p/word2vec/). It acts on a 0 to 1 scale, with higher values indicating a smaller semantic distance between words.

Standard deviations are given in parentheses.

5.3.2.4 Procedure and design

Participants were asked to name pictures in Polish and English depending on their preference. All instructions were provided on the computer screen in Polish. The task started with a familiarization phase, followed by the experimental phase.

During the familiarization phase, participants were asked to familiarize themselves with the 16 pictures and their corresponding names in Polish and English. The participants were told to remember the names to use them later in the experimental phase. The pictures, along with their names, were presented one by one on the screen. The participants could move through the familiarization phase at their own speed and repeat it until they felt comfortable enough to move on to the experimental phase.

During the experimental phase, participants were told to use the picture names from the familiarization phase but were free to choose the language (Polish or English). The following instructions were provided prior to the experimental phase: 'You are free to switch between languages whenever you want. Try to use the word that comes to mind first, but don't use the same language throughout the whole task'. All instructions were given in Polish.

The experimental phase consisted of 384 trials, distributed evenly across eight blocks (48 trials each) with short self-paced breaks in between. A fixed order of language preference stimuli sets (Polish, English and ambiguous preference) was used within each block to allow the investigation of how different stimuli sets affect language choice. There was a maximum of three consecutive trials of the same language preference stimuli set. However, the pictures belonging to the same stimuli set were randomly selected on each trial, e.g., on a single Polish preference trial, one of the four pictures belonging to the PL set was randomly chosen. In total, there were 96 trials with Polish preference including 4 pictures, 96 trials with English preference including 4 pictures, and 192 trials with ambiguous language preference including 8 pictures.

Each trial started with the presentation of a jittered fixation cross (300, 400, or 500 ms), followed by the presentation of the picture. The pictures were displayed on the screen for 2500 ms, regardless of when a response was given. Afterward, a white screen was presented for 300 ms. The experimental phase was preceded by 16 randomized practice trials with the same stimuli and design as the experimental trials.

5.3.2.5 Analysis

Trials without a response or with incorrect or unidentifiable responses were discarded (6.07% of the total data). Trials after an incorrect response, as well as the first trial of each block, were excluded from the analysis (4.52% of the total data), as it was not possible to determine if they were repetition or switching trials. The remaining trials were coded as repeat trials (i.e., naming the current trial in the same language as the previous trial) or switch trials (i.e., naming the current trial in a different language than the previous trial). Reaction times (RTs) were derived using Chronset software (Roux et al., 2017). Outlier RTs (i.e., RTs below 300 ms, above 5000 ms, or 3 SD below/above a mean) were discarded from the analysis (1.48% of the total data).

To investigate the voluntary language switching pattern, two within-subject factors, i.e., Language (Polish vs. English) and Trial type (repeat vs. switch), were examined using two-way repeated-measures ANOVA with RTs as the dependent variable. We expected Trial type to be significant, demonstrating that our participants were slower during switch trials compared to repeat trials (switch cost effect). A significant interaction between Language and Trial type would indicate switch cost asymmetry. We predicted that switch trials in Polish show larger RTs than switch trials in English. We performed a post hoc *t*-test comparing the switch costs (index created by subtracting RTs of repeat trials from RTs of switch trials) in both languages to test this prediction.

To analyze the impact of the language preference stimuli sets on language choice, the within-subject factors Language preference stimuli set (Polish vs. English vs. ambiguous preference) and Language of response (Polish vs. English) were examined using a two-way repeated-measures ANOVA with RTs as the dependent variable. A significant interaction between Language preference and Language would indicate that language preference impacts language choice. We predicted that pictures from the Polish and English preference sets were more likely to be named in Polish or English, respectively. Additionally, we performed posthoc *t*-tests comparing RTs between the three stimuli sets in Polish and English. We predicted that pictures from the Polish preference stimuli set would be named faster in Polish compared to English, and pictures from the English preference stimuli set would be named faster in English compared to Polish. Stimuli from the ambiguous preference stimuli set, we predicted, would show no significant difference between languages.

5.3.2.6 Results

The analysis of voluntary language switching patterns revealed a main effect of Language, *F*(1, 49) = 27.78; *Mse* = 86874; p < .01, with participants naming pictures significantly slower in Polish (*M* = 900.30 ms, *SD* = 141.63) than in English (*M* = 858.62 ms, *SD* = 124.06). The effect of Trial type was also significant, *F*(1, 49) = 60.83; *Mse* = 73179; *p* < .01, with switch trials (*M* = 898.59 ms, *SD* = 138.98) being significantly slower than repeat trials (*M* = 860.33 ms, *SD* = 127.57). These main effects were not qualified by an interaction between Language and Trial type, *F*(1, 49) = 1.28; *Mse* = 1142; *p* = .26.

The analysis of language preference revealed a main effect of Language, *F*(1, 48) = 24.64; *Mse* = 130150; *p* < .01, with participants naming pictures significantly slower in Polish (*M* = 903.11 ms, *SD* = 136.54) than in English (*M* = 861.03 ms, *SD* = 121.31). The effect of Language preference was also significant, *F*(1, 96) = 9.17; *Mse* = 12461; *p* < .01. The follow-up pairwise t-tests with Bonferroni correction further showed that stimuli with ambiguous preference (*M* = 869.13 ms, *SD* = 128.01) were significantly faster named than stimuli with Polish preference (*M* = 887.26 ms, *SD* = 128.18, *p* < .001) and English preference (*M* = 889.81 ms, *SD* = 135.95, *p* < .001). At the same time, there was no significant difference between

Polish and English preference stimuli sets (*p* = 1). The interaction between Language and Language preference was not significant, *F*(1, 96) = 1.7; *Mse* = 2300; *p* = .19, suggesting that the selected stimuli sets had no impact on the language choice during voluntary language switching.

5.3.3 Methodology Main experiment

5.3.3.1 Participants

We recruited 103 Polish-English bilinguals for this study. The number of required participants was determined by a previous power analysis using data of a similar training study (Timmer et al., 2019) and the SIMR package for R (P. Green & MacLeod, 2016). Only right-handed individuals between the ages of 20 to 35 years living in Poland were recruited. The participants were native Polish speakers with scores of at least 60% on the English LexTALE (Lemhöfer & Broersma, 2012) and the General English Test (by Cambridge Assessment: https://www.cambridgeenglish.org/test-your-english/general-english/). The participants were also required to have normal or corrected-to-normal vision without color-blindness and no history of neurological impairments or language disorders.

All participants completed a language background questionnaire based on Li et al. (2014) and Marian Viorica et al. (2007). Prior to the start of the experiment, they were assigned to one of the two experimental groups: the forced language switching group or the voluntary language switching group. To ensure an even distribution of participants over the groups, we used individual matching: each participant of the forced language switching group was matched with a similar participant of the voluntary language switching group based on age, gender, age of L2 acquisition, L2 proficiency (based on LexTALE and the General English Test) and their socioeconomic status (Adler et al., 2000). Based on the participants included in the final analyses, the groups were statistically equivalent on age: $t(83) = 1.85$, gender: $t(83) = - 0.69$, age of L2 acquisition: $t(83) = -0.65$, L2 proficiency: $t(83) = 0.42$ and socioeconomic status: $t(83) = -0.35$.

Data from eighteen participants were eliminated due to technical problems in the experimental task and EEG recording (forced language switching: n = 6; voluntary language switching n = 4) or high error rates (30% incorrect on the mixed blocks of the nonverbal task; forced language switching: n = 1; voluntary language switching n = 2). Five bilinguals of the voluntary language switching group did not switch tasks enough in the voluntary language switching block to produce more than 20 trials in the four conditions of interest (i.e., repeat and switch trials in Polish and English) and were excluded from the analyses. Following these exclusions, 85 bilinguals contributed data to the analysis, 44 in the forced language switching group and 41 in the voluntary language switching group. Table 3.2 presents an overview of background measures and demographics. In addition to English, participants of both groups spoke a variety of languages, which includes German (n = 6), Spanish (n = 5), French (n = 5), Italian (n = 3), Russian (n = 2) and Swedish (n = 2). Single participants also spoke Arabic, Dutch, Japanese and Portuguese.

Table 3.2

Demographic information and language experience of participants included in the final analysis.

	Forced language switching group		Voluntary language switching group		t-test	
N	44 (8 female)		41 (10 female)			
Participant’s Age (years)	23.66 (2.99)		22.51 (2.72)		*t*(83) = 1.85, *p* = .07	
SES (1 – 10)	5.80 (1.34)		5.90 (1.46)		*t*(83) = -0.35, *p* = .73	
Self-assessed language experience (1-10)	L1	L2	L1	L2	L1	L2
Self-rated proficiency	8.99 (0.05)	7.64 (0.88)	8.99 (0.04)	7.48 (0.71)	*t*(83) = -0.52, *p* = .60	*t*(83) = -0.89, *p* = .38
Speaking	9.00 (0.00)	7.18 (1.19)	8.98 (0.16)	7.15 (0.88)	*t*(83) = -1.04, *p* = .30	*t*(83) = -0.16, *p* = .88
Writing	8.95 (0.21)	7.36 (0.99)	9.00 (0.00)	7.12 (1.08)	*t*(83) = -0.28, *p* = .78	*t*(83) = 1.08, *p* = .28
Listening	9.00 (0.00)	7.89 (0.95)	9.00 (0.00)	7.68 (0.79)	-	*t*(83) = -1.07, *p* = .29
Reading	9.00 (0.00)	8.11 (0.95)	9.00 (0.00)	7.98 (0.76)	-	*t*(83) = -0.74, *p* = .46
Age of L2 acquisition (years)	-	5.86 (2.03)	-	6.15 (1.98)	*t*(83) = -0.0.65, *p* = .52	

Frequency of intrasentential code-switching	0.68 (0.37)		0.73 (0.39)		*t*(83) = -0.60, *p* = .55
Objective L2 proficiency measures					
LexTALE (mean accuracy in %)	-	76.36 (10.43)	-	74.59 (10.14)	*t*(83) = 0.79, *p* = .43
General English Test (mean accuracy in %)	-	87.36 (9.00)	-	87.51 (9.85)	*t*(83) = -0.07, *p* = .94

Note. The first part of the table describes the demographic information of the final forced language switching group and the final voluntary language switching group. The rows display (1) the number of participants with the number of women in brackets, (2) age (in years) and (3) socio-economic status on a 1 to 8 scale based on Adler et al. (2000), The second part of the table summarizes the self-assessed language experience based on a questionnaire. The (4) 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 (5) age of L2 acquisition is in years. The (6) frequency of language switching within sentences is presented on a scale from 1 to 9, where 1= "I never switch languages within sentences" and 9 = "I always switch languages within sentences". The objective L2 proficiency measures in English, (7) the accuracy in the LexTALE task and (8) the accuracy in the General English Test, are presented in percentages. Standard deviations are given in parentheses.

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

5.3.3.2 General procedure

Experimental procedures for data collection and analysis were approved by the Institutional Review Board of Jagiellonian University. The participants signed an informed consent form prior to the experiment and were paid PLN 50 for their participation in the study. Participants were tested individually in a quiet room with dimmed lights and seated approximately 80 cm from the computer screen. We recorded both their behavioral data and their brain activity using electroencephalography (EEG).

The experiment consisted of a pre-test, language switching training, and a post-test. During pre-test and post-test, all participants completed a non-linguistic task switching paradigm (TS). During the language training, the participants completed a picture-naming task which differed in the task requirements between the two participant groups (forced language switching group or the voluntary language switching group). After the pre-test and the language switching training, resting state EEG was recorded for 5 minutes. Additionally, the participants rated the difficulty of the language switching training on a cognitive effort questionnaire (Leppink et al., 2013).

5.3.3.3 Procedure and design

5.3.3.3.1 Pre- and post-test: non-linguistic task switching

We used a color/shape task switching task based on Timmer, Calabria, and Costa (2019). Participants saw colored stimuli created from four shapes (square, triangle, circle, and oval) and four colors (yellow, green, white, and grey). They made decisions about the color (colored or greyscale) or shape (corners or rounded) by pressing the ‘Z’ or ‘M’ keyboard button for each stimulus with their left and right index fingers, respectively. To indicate which decision type (color or shape) should be performed, a cue (solid or dashed outline, respectively) appeared around the stimulus. The assignment of cues to decision types and responses to corresponding buttons was counterbalanced across participants, as was the assignment of keys (‘Z’ or ‘M’) to a response (color or greyscale and corners or rounded). Only a subset of the stimuli was presented to each participant to allow the determination of the accuracy of each response. For example, a yellow triangle was not presented if a participant was responding with ‘Z’ for colored (color task) and ‘Z’ for corners (shape task). In this case, the correct responses for the color and shape task converged on the same keys, and no judgment of accuracy could have been made.

The switching task started with two pure blocks during which only one decision type was performed in each (color or shape, order counterbalanced across participants). This was followed by four mixed blocks that included both types of rules. The task ended with again two pure blocks. Participants were able to take short breaks after every block. Each pure block consisted of 32 trials (128 pure trials in total for each participant). The mixed condition consisted of 4 blocks of 48 trials each (192 mixed block trials in total). The repeat-to-switch ratio was 50%/50%, with a total of 96 repeat and 96 switch trials for each participant. Likewise, the number of color and shape trials was kept equal to 96 color task trials and 96 shape task trials by using a fixed order of cues. An additional trial was added at the beginning of each block (both the pure and mixed blocks), as the first trial cannot be coded as either a switch or repeat trial.

The pure blocks were preceded by eight practice trials and the mixed blocks by 16 practice trials. Participants were given the opportunity to repeat the practice trials until they felt comfortable enough to move on to the experimental blocks. Each trial was initiated by a fixation cross with a random duration of 400, 500, or 600 ms, followed by the presentation of the target picture together with the cue. The picture remained on the screen until a response was given or after 2500 ms had passed. Only during the practice trials was feedback provided ("correct" or "incorrect") for 750 ms. All trials ended with the presentation of a blank screen for 500 ms.

5.3.3.3.2 Language switching training

The participants completed a picture-naming task with language switching between Polish and English. Participants were assigned to one of the two types of training: forced language switching training or voluntary language switching training. The two training types only differed in the use of language cues. In the forced language switching group, a cue accompanying the presentation of a picture indicated the language in which a picture should be named (forced picture-naming task). In the voluntary language switching group, the same cue was presented throughout (voluntary picture-naming task). Stimuli were the identical 16 colored pictures of Pilot 2 (for description, see section 5.3.2.3.). Each language switching training session started with a familiarization phase, followed by the main training phase.

The language switching training was divided into four blocks. Participants were able to take short breaks after every block. Each block consisted of 160 trials, with each of the 16 pictures presented ten times for a total of 640 trials for each participant.

In the forced picture-naming task, a solid or dashed outline around each picture was used to indicate naming in Polish or English respectively. A fixed cue order was used to hold the repeat-to-switch ratio at 50%, with a total of 80 repeat and 80 switch trials for each participant. Four different cue orders were created and counterbalanced across blocks. In the voluntary picture-naming task, the same outline (solid or dashed) was used for all trials, allowing a free language choice. The assignment of cue outline was counterbalanced across participants for both groups.

For both voluntary and forced picture-naming tasks, each trial was initiated by a fixation cross with a random duration of 400, 500 or 600 ms followed by the presentation of the target picture together with the cue. The picture remained on the screen until a response was given or after 2500 ms had passed. The voice onset was detected with a microphone and the SR-box (Schneider, 1995). All trials ended with the presentation of a blank screen for 500 ms. The language switching training session was preceded by a familiarization phase as applied in Pilot 2 (for a description, see section 5.3.2.4.).

5.3.3.3.3 Apparatus and data acquisition

The EEG signal was continuously recorded from 64 Ag/AgCl active electrodes (with preamplifiers) using the BioSemi ActiveTwo system. The electrodes were secured in an elastic cap according to the extended 10–20 international electrode placement system. The zero-reference principal voltage values (each site was quantified relative to the DRL and CMS loop) were digitized at a sampling rate of 512 Hz. Two electrodes above and below the left eye recorded the eye blinks; another two electrodes (placed at the external canthi of each eye) recorded horizontal eye movements. Two additional electrodes were placed on the left and right mastoids as references. Finally, another set of two electrodes was placed on the lower chest to record heart rate (EKG).

The data was pre-processed in Matlab using EEGlab (Delorme & Makeig, 2004) and ERPlab (Lopez-Calderon & Luck, 2014). All channels were re-referenced to the average of the left and right mastoids. The EEG signal was down sampled to 256 Hz and then filtered with a high-pass filter of 0.1 Hz (12 dB), a low-pass filter of 40 Hz (24 dB), and a notch filter. Ocular artifacts were corrected using Independent Component Analysis (Delorme et al., 2007; Jung et al., 2000). Afterwards, stimulus-locked epochs of -200 to 700 ms were extracted from the continuous EEG signal and aligned to the pre-stimulus baseline from -200 ms to 0 ms. The segments were checked separately for each trial type in the TS (i.e., switch and repeat, color and shape).

The selection of electrode clusters and time windows for the targeted ERP components was guided by previous research using similar experimental tasks. For both task switching and task mixing we focused on the parietal P3 (Gajewski & Falkenstein, 2011; Jost, Mayr, & Rösler, 2008; Timmer, Grundy, & Bialystok, 2017; Wong et al., 2018). Consequently, we analysed a pronounced positivity around the parietal region (CPz, Pz, CP1, CP2) in the 300-450 ms time window after target onset. Mean voltage amplitudes in the electrode clusters and time windows for each trial were used for statistical analysis.

5.3.3.3.4 Analysis

The data were analyzed using linear mixed-effects models (LME). All models were fitted using the lme4 package in R (version 1.1-13; R Core Team, 2019). In the analyses, we focused on two dimensions of cognitive efficiency: non-linguistic *switching* and *mixing* as measured in the TS task. We used RTs and the P3 mean amplitude as dependent variables for both dimensions and specified one model, as presented below. Due to the right-skewed distribution of reaction times, they were transformed using reciprocal transformation (-1000/reaction time).

The models of each dimension had an identical structure. Only a single variable, the experimental variable hereafter, was changed between the two models. The two experimental variables for the two dimensions were: Trial type (switch or repeat) for switch costs*,* and Block type (pure trials from single block or repeat trials from mixed block) for mixing costs.

Every model included Participants and Items as crossed random effects. As fixed effects, we included Test phase (pre- or post-test), Group (forced or voluntary language switching), and their interaction. Each model also included the experimental variable, as well as a two-way interaction between the experimental variable and Test phase and a three-way interaction between the experimental variable, Test phase and Group. The maximal models also included by-Participant and by-Item random intercepts and random slopes for Test phase, the Experimental variable, and their interaction for the Participant intercept and random slopes for Test phase, Group, Trial Type, and their interactions for the Item intercept. All independent variables were coded using a difference contrast as recommended by Schad and colleagues (2020). We fitted first the maximal model (Barr, Levy, Scheepers, & Tily, 2013). When 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, Kliegl, Vasishth, and Baayen (2018). Absolute *t*-values greater than the conventional level of two were considered significant.

For both models, we predicted that Test phase would be significant with shorter RTs for the post-test compared to the pre-test, showing the effect of repeating the task. However, the two models differed in their predictions for the two investigated dimensions of cognitive efficiency.

The switch cost model served to test how language switching training affected switch costs in the TS. We predicted a main effect of Trial type with shorter RTs and an increased P3 amplitude for repeat compared to switch trials, confirming our task design (P. D. Gajewski & Falkenstein, 2011; Timmer et al., 2017). The interaction between Test phase and Trial type was predicted to be significant and show the effect of repetition of the TS with a reduced effect of Trial type for the post-test. Following the ACH, we predicted a significant three-way interaction between Test phase, Group, and Trial type, which would indicate the differences in the efficiency of task switching related to language switching training.

The mixing cost model served to test how language switching training affected mixing costs in the TS. We predicted that Block type would be significant with shorter RTs and an increased P3 amplitude for pure trials from the single blocks compared to repeat trials from the mixed blocks, confirming our task design (Timmer et al., 2017). The interaction between Test phase and Block type would be significant and show the effect of repetition of the TS with a reduced effect of Block type for the post-test. We expected the three-way interaction between Test phase, Group, and Block type to not be significant as global control mechanisms receive no additional activation during forced language switching.

Post-hoc tests for the three-way interactions were done using the emmeans package for R (Lenth, 2021) for only the post-test data.

5.4 Results

5.4.1 Behavioral results

5.4.1.1 Language switching training

We conducted two versions of the language switching task. In the forced language switching task, the participant's language use was controlled by the presented cue with 50% switch trials and 50% English trials. In the voluntary language switching task, participants could choose their language based on their own volition. When analyzing the accurately named trials, the switching frequency in both groups showed that participants, on average, switched on 49.0% (*SD* = 0.8, range 46.2% - 50.9%) of the trials in the forced language switching group and 38.5% (*SD* = 8.5, range 22.9% - 60.1%) of the trials in the voluntary language switching group. The forced language switching group named 51.7% (*SD* = 1.0, range 50.3% - 56.1%) of trials in English, and the voluntary language switching group 55.9% (*SD* = 5.6, range 47.6% - 76.7%). The deviation by the forced language switching group from the targeted 50% switch trials and 50% English trials was due to erroneous answers by the participants.

The behavioral data for the language switching training are presented in Table 3.2. We found a significant effect of Group with the forced language switching group (*M* = 990.49 ms, *SD* = 263.96) naming pictures significantly slower than the voluntary language switching group (*M* = 792.90 ms, *SD* = 198.50). In addition, there was a significant Trial type effect with switch trials (forced: *M* = 1031.09 ms, *SD* = 268.14; voluntary: *M* = 812.09 ms, *SD* = 204.92) showing longer naming latencies than repeat trials (forced: *M* = 950.35 ms, *SD* = 253.45: voluntary: *M* = 783.99 ms, *SD* = 194.81). Surprisingly, we found reversed language asymmetry, with Polish switch costs (*M* = 42.77 ms, *SD* = 62.97) being smaller than English ones (*M* = 53.04 ms, *SD* = 49.43). The switch costs also differed significantly between groups, with the forced language switching group (*M* = 75.65 ms, *SD* = 55.44) showing larger switch costs than the voluntary language switching group (*M* = 19.07 ms, *SD* = 41.87). For the used languages, we also found slower naming latencies for Polish trials (forced: *M* = 1037.71 ms, *SD* = 274.04; voluntary: *M*

= 832.26 ms, *SD* = 222.58) than English trials (forced: *M* = 948.59 ms, *SD* = 247.25; voluntary: *M* = 768.93 *SD* = 178.07). Group also modulated the Language effect as a significant Language Group interaction with a smaller Language effect for the voluntary language switching group than the forced language switching group.

Table 3.2

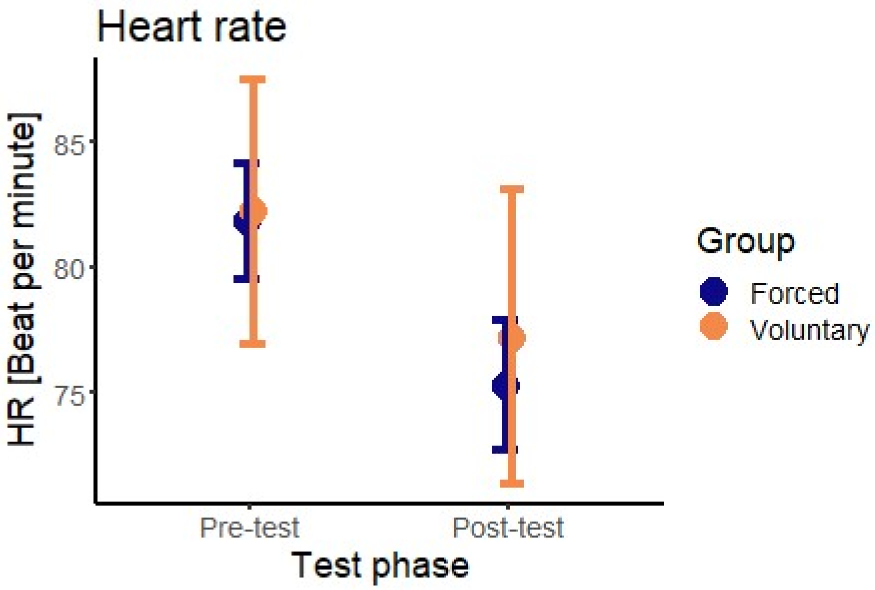
Results of the linear mixed model for language switching training with both groups.

	Estimate	SE	t	p	by-Item SD	by-Participant SD
Intercept	-1.18	0.02	-56.621	< 0.001***	0.00	0.02
Trial type	-0.06	0.00	-24.529	< 0.001***		
Language	-0.08	0.02	-4.323	< 0.001***	0.01	0.00
Group	-0.23	0.03	-7.369	< 0.001***		-
Trial type:Language	-0.01	0.00	-3.136	< 0.01**		
Trial type:Group	0.06	0.00	12.158	< 0.001***		-
Language:Group	0.04	0.01	2.912	< 0.01**		-
Trial type:Language:Group	0.01	0.01	0.872			-

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

To assess the difficulty of the language switching tasks, we employed two measures: a self-rated effort questionnaire (Leppink et al., 2013) and heart rate measured during the pre- and post-test. The analysis of the self-rated effort revealed no significant effect of Group, *t*(83) = 1.54, *p* = .13, with both groups giving similar ratings to their respective language switching training (forced: *M* = 4.05, *SD* = 0.39; voluntary: *M* = 3.88, *SD* = 0.65). The analysis of the heart rate demonstrated a main effect of Test phase (Table 3.3), with participants having a lower heart rate in the post-test (*M* = 76.20 bpm, *SD* = 11.12) than in the pre-test (*M* = 82.02 bpm, *SD* = 11.28). However, there was no effect of Group (forced: *M* = 78.54 bpm, *SD* = 10.51; voluntary: *M* = 79.72 bpm, *SD* = 12.58). Together these results indicate that no additional stress was placed on the forced language switching group compared to the voluntary language switching group. The slow-down of the heart in the post-test also showed that the participants adapted to the experimental setup and were calmer during the post-test.

Figure 3.1



Note. Average heart rates of both participant groups in both experiment test phases. Error bars represent standard deviation.

Table 3.3

Results of the linear mixed model for heart rate with both groups.

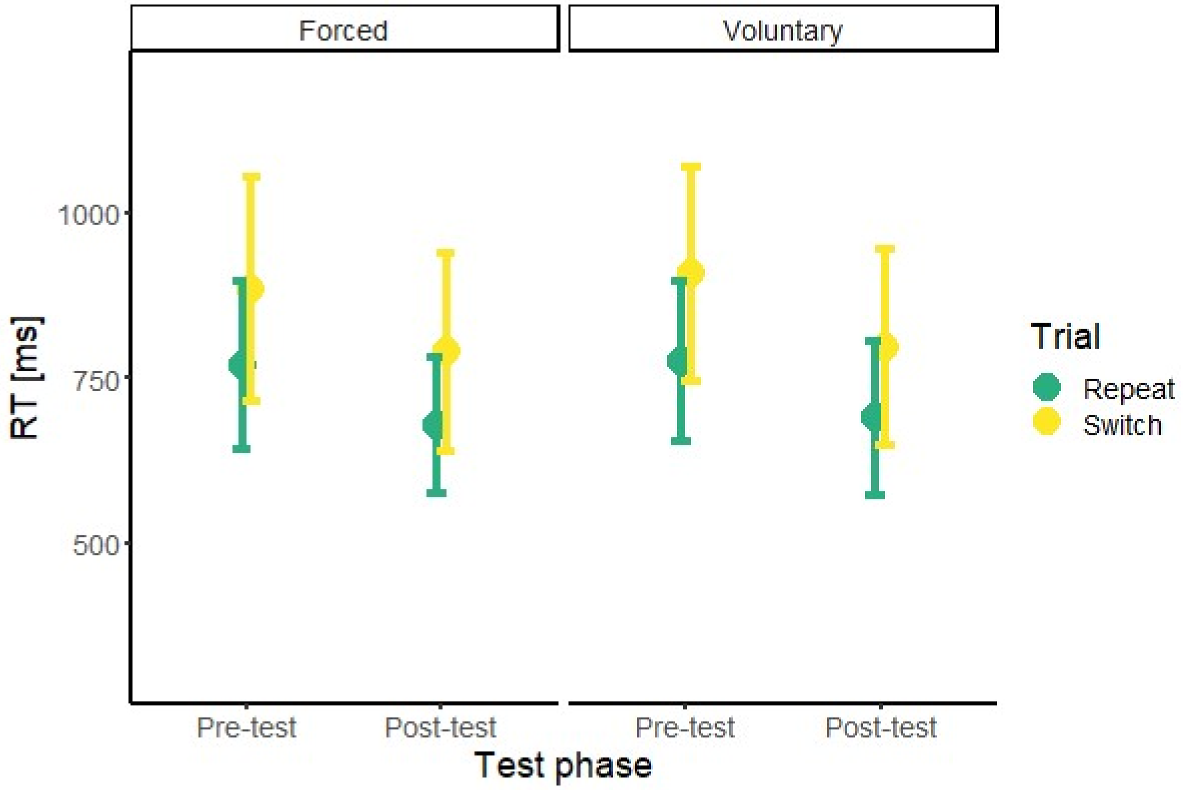
	Estimate	SE	t	p	by-Item SD
Intercept	79.13	1.19	66.75***	< 0.001***	10.59
Test phase	-5.79	0.58	-10.03***	< 0.001***	-
Group	1.18	2.37	0.50	0.62	-
Test phase:Group	1.54	1.15	1.34	0.19	-

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

5.4.1.2 Pre- and post-test analysis - switch costs

The behavioral results of the non-linguistic switching task are presented in Table 3.4. For the behavioral switch cost model (Table 3.6), we found a significant effect of Test phase, with participants being overall faster during the post-test. Additionally, Trial type was significant, with switch trials being slower than repeat trials (Figure 3.2). However, we found no evidence for an overall group difference or a difference in the training effect between groups.

Figure 3.2



Note. Average reaction times for switch costs in the non-linguistic task switching of both participant groups in both experiment test phases. Error bars represent standard deviation.

Table 3.4

Mean response latencies in ms (and standard deviation) and mean amplitude of the P3 component (and standard deviation) for the non-linguistic task switching.

		Pre-test			Post-test		
		Pure blocks	Mixed blocks		Pure blocks	Mixed blocks	
		Pure	Repeat	Switch	Pure	Repeat	Switch
RT (ms)	Forced language switching group	531.96 (176.28)	774.09 (213.10)	889.19 (219.16)	489.74 (155.28)	681.21 (206.19)	787.07 (220.54)
	Voluntary language switching group	531.08 (180.05)	778.59 (229.49)	907.93 (236.55)	483.60 (167.79)	688.66 (222.40)	791.02 (229.79)
Voltage (mV)	Forced language switching group	5.77 (15.70)	5.29 (15.21)	5.51 (15.45)	6.39 (17.90)	5.27 (17.25)	5.68 (16.99)
	Voluntary language switching group	5.79 (17.29)	4.94 (16.15)	5.62 (15.80)	6.69 (17.77)	4.48 (17.19)	5.01 (17.13)

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

Table 3.5

Post hoc analyses for differences in trial type and block type for the post-test test phase of the non-linguistic task switching.

		Switch trials Estimated marginal means (95% CI)	Repeat trials Estimated marginal means (95% CI)	Mean difference estimate	Standard Error (SE)	*P* value
Transformed RT (ms)	Forced language switching group	-1.61 (-1.68, -1.54)	-1.38 (-1.45, -1.31)	-0.23	0.01	< .001***
	Voluntary language switching group	-1.59 (-1.66, -1.52)	-1.36 (-1.44, -1.29)	-0.23	0.01	< .001***
Voltage (mV)	Forced language switching group	5.27 (4.44, 6.09)	5.65 (4.80, 6.50)	-0.38	0.43	.37
	Voluntary language switching group	4.46 (3.62, 5.31)	5.01 (4.14, 5.88)	-0.55	0.44	.21
		Pure trials Estimated marginal means (95% CI)	Repeat trials Estimated marginal means (95% CI)	Mean difference estimate	Standard Error (SE)	*P* value
Transformed RT (ms)	Forced language switching group	-2.21 (-2.29, -2.13)	-1.60 (-1.69, -1.52)	-0.61	0.01	< .001***
	Voluntary language switching group	-2.26 (-2.34, -2.18)	-1.59 (-1.67, -1.50)	-0.68	0.01	< .001***
Voltage (mV)	Forced language switching group	6.39 (5.61, 7.17)	5.29 (4.42, 6.15)	1.1	0.39	< .01**
	Voluntary language switching group	6.67 (5.86, 7.48)	4.47 (3.58, 5.35)	2.2	0.40	< .001***

Note. Post hoc tests were performed with only data from the post-test. Separate models were run for switch costs and mixing costs. Reaction times are reciprocally transformed.

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

Table 3.6

Results of the linear mixed model for reaction times for switch costs in non-linguistic task switching paradigm for both groups.

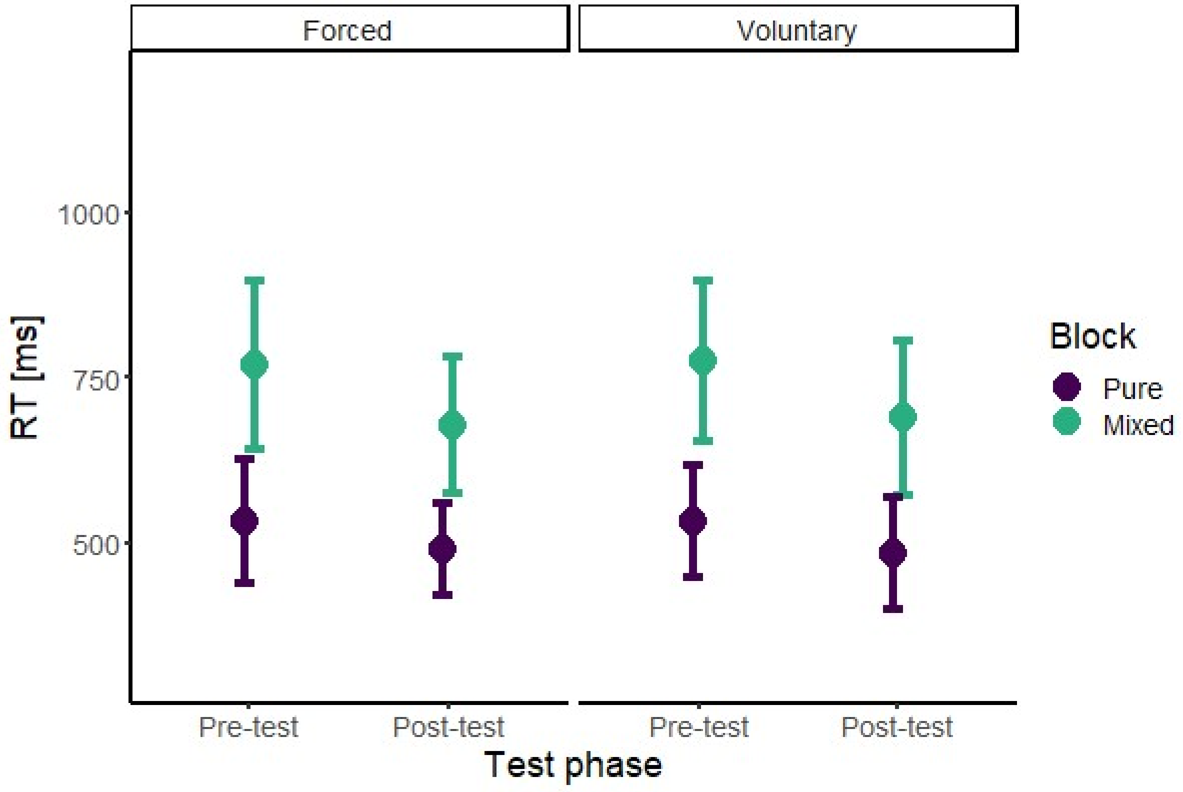
	Estimate	SE	t	p	by-Item SD	by-Participant SD
Intercept	-1.39	0.03	-55.62	<0.001***	0.03	0.22
Test phase	-0.18	0.01	-17.91	<0.001***		0.08
Group	0.02	0.05	0.38	0.70		-
Trial type	0.22	0.01	16.65	<0.001***		0.11
Test phase:Group	-0.01	0.02	-0.34	0.73		-
Test phase:Trial type	0.02	0.01	2.57	0.01*		
Pre-test:Group:Trial type	0.02	0.03	0.79	0.43		-
Post-test:Group:Trial type	<0.01	0.03	-0.11	0.92		-

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

5.4.1.3 Pre- and post-test analysis - mixing costs

The mixing cost model for the behavioral data found a significant effect for Test phase (Table 3.7), demonstrating that the participants improved between the two experimental test phases. Additionally, the effect of Block type was significant with pure trials from the single block named faster than repeat trials from the mixed blocks (Figure 3.3). There was also a significant interaction between Test phase and Block type, with reduced mixing costs in the post-test compared to the pre-test. Importantly, we found a significant three-way interaction of Test phase, Block Type, and Group, with the forced language switching group showing a smaller mixing effect during the post-test test phase than the voluntary language switching group (Table 3.5).

Figure 3.3



Note. Average reaction times for mixing costs in the non-linguistic task switching of both participant groups in both experiment test phases. Error bars represent standard deviation.

Table 3.7

Results of the linear mixed model for reaction times for mixing costs in non-linguistic task switching with both groups.

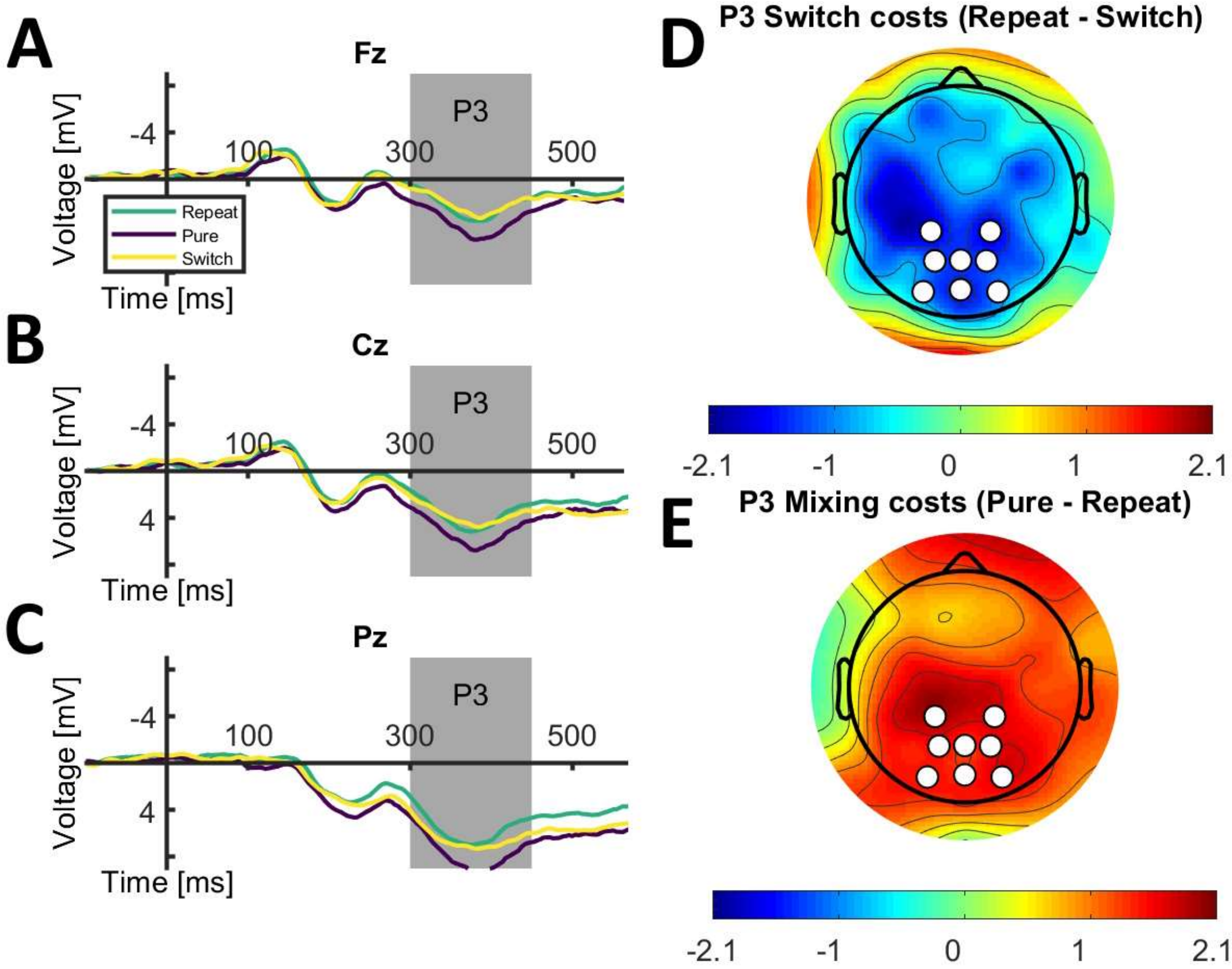
	Estimate	SE	t	p	by-Item SD	by-Participant SD
Intercept	-1.82	0.03	-62.22	<0.001***	0.04	0.25
Test phase	-0.19	0.01	-33.43	<0.001***		
Group	-0.01	0.06	-0.13	0.90		-
Block type	0.65	0.01	116.18	<0.001***		
Test phase:Group	-0.02	0.01	-1.38	0.17		-
Test phase:Block type	-0.03	0.01	-2.48	0.01*		
Pre-test:Group:Block type	0.02	0.02	1.13	0.26		-
Post-test:Group:Block type	0.07	0.02	4.45	<0.001***		-

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

5.4.2 EEG results

Figure 3.4 shows an overview of the recorded EEG activity during the non-linguistic task switching. We found a generally parietal distribution for the target-locked P3 component for both switch and mixing costs.

Figure 3.4

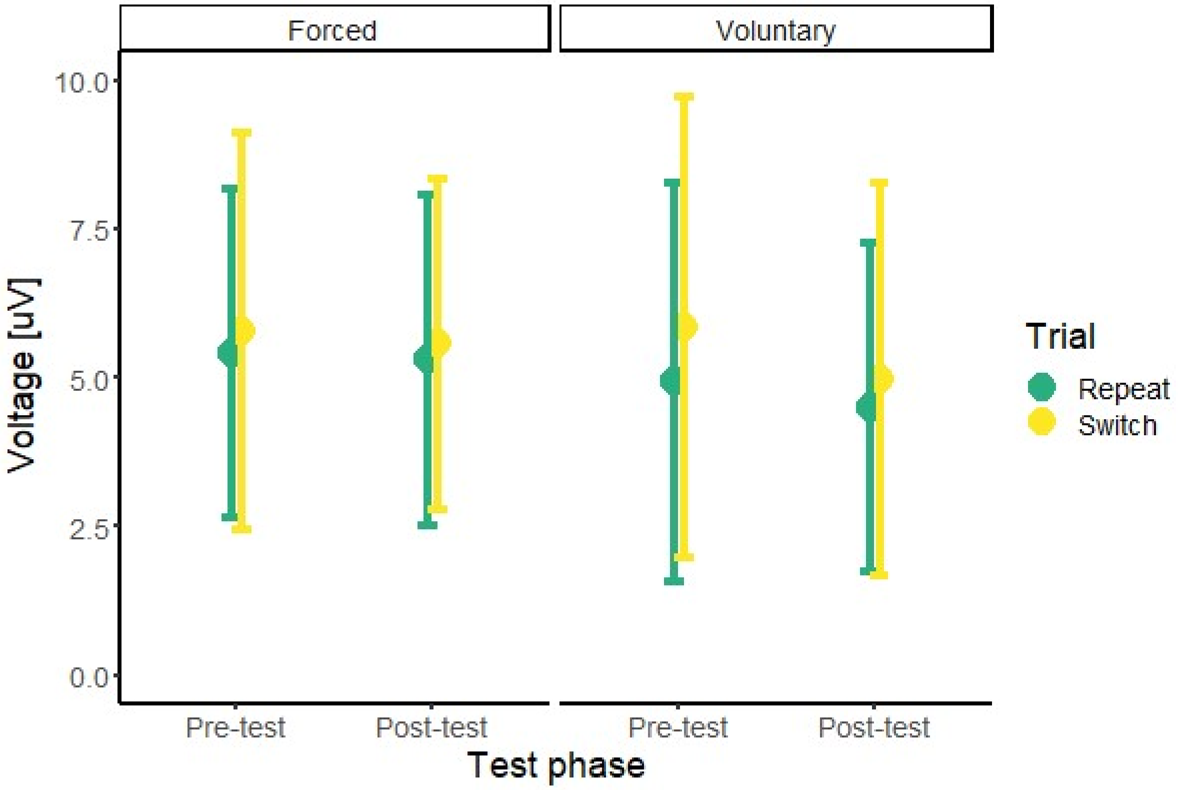


Note. Mean amplitudes for the three trial types at the midline electrodes Fz (A), Cz (B), and Pz (C) and scalp topography maps for the P3 component for switch costs (D) and mixing costs €. Mean amplitudes were averaged over pre- and post-test, as well as group. Selected electrode clusters used for calculating mean amplitudes are marked in white.

5.4.2.1 Pre- and post-test analysis - switch costs

In our analysis of the P3 mean amplitude, the switch costs model revealed a significant effect of Trial type, with switch trials showing a more positive amplitude than repeat trials (Figure 3.5). No other significant effects were observed. The results of the P3 response to the non-linguistic task switching are found in Table 3.8.

Figure 3.5



Note. Mean P3 amplitudes for switch costs in the non-linguistic task switching of both participant groups in both experiment test phases. Error bars represent standard deviation.

Table 3.8

Results of the linear mixed model for P3 mean amplitude for switch costs in non-linguistic task switching with both groups.

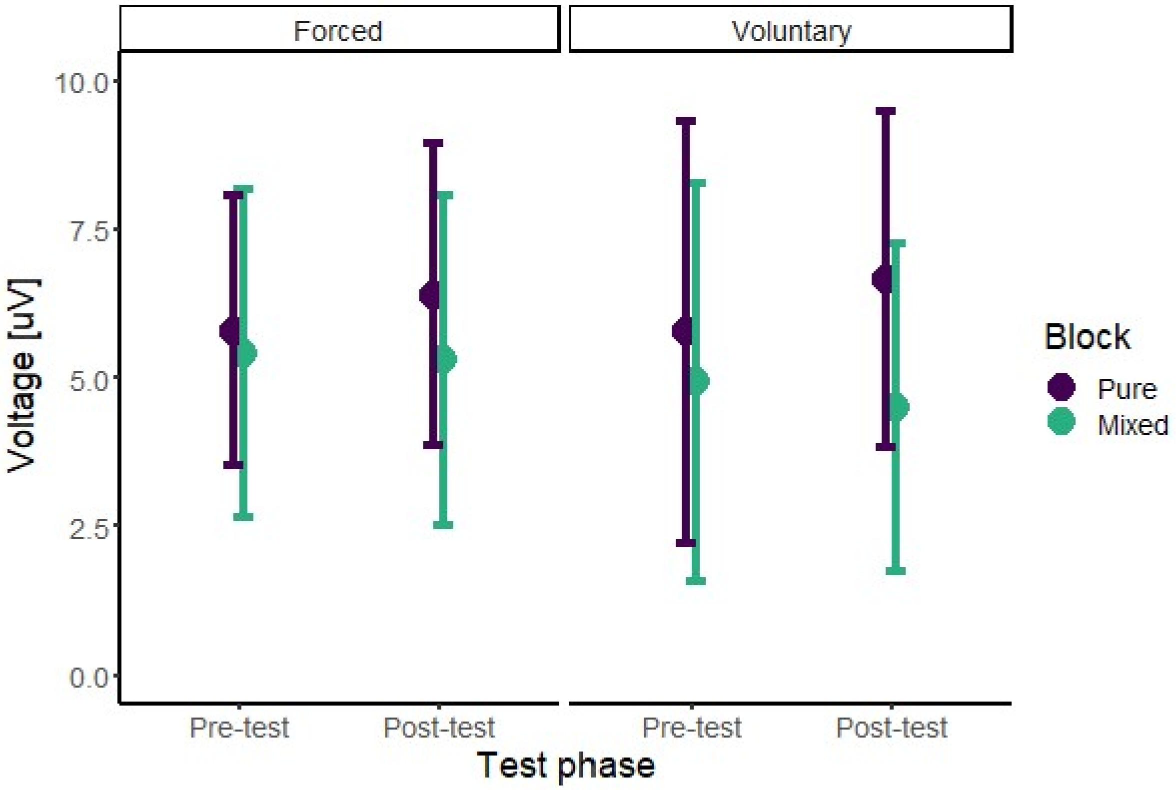
	Estimate	SE	t	p	by-Item SD	by-Participant SD
Intercept	5.26	0.25	21.22	<0.001***		2.01
Test phase	-0.32	0.23	-1.36	0.17		
Group	-0.44	0.50	-0.89	0.38		-
Trial type	0.52	0.23	2.23	0.03*		
Test phase:Group	-0.54	0.47	-1.16	0.25		-
Test phase:Trial type	-0.09	0.46	-0.19	0.85		
Pre-test:Group:Trial type	0.56	0.72	0.78	0.44		-
Post-test:Group:Trial type	0.19	0.59	0.32	0.75		-

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

5.4.2.2 Pre- and post-test analysis - mixing costs

The results of the P3 response to mixing costs in non-linguistic task switching were analyzed with the mixing cost model (Table 3.9). We found a significant effect of Block Type on the P3 amplitude, with pure trials having a more positive amplitude than repeat trials from the mixed blocks (Table 3.9). The interaction between Block Type and Test phase was also significant, with a larger difference between pure and repeat trials during the post-test. We also found a significant three-way interaction between Test phase, Block Type, and Group, with the voluntary language switching group having a significantly larger Block type effect during the post-test than the forced language switching group (Table 3.5).

Figure 3.6



Note. Mean P3 amplitudes for mixing costs in the non-linguistic task switching of both participant groups in both experiment test phases. Error bars represent standard deviation.

Table 3.9

Results of the linear mixed model for P3 mean amplitude for mixing costs in non-linguistic task switching with both groups.

	Estimate	SE	t	p	by-Item SD	by-Participant SD
Intercept	5.59	0.25	22.03	<0.001***	0.01	2.15
Test phase	0.26	0.20	1.31	0.19		
Group	-0.27	0.51	-0.53	0.60		-
Block type	-1.15	0.20	-5.87	<0.001***		
Test phase:Group	-0.02	0.39	-0.04	0.97		-
Test phase:Block type	-1.00	0.39	-2.55	0.01*		
Pre-test:Group:Block type	-0.50	0.58	-0.88	0.38		-
Post-test:Group:Block type	-1.11	0.53	-2.10	0.04*		-

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

5.4.3 Associations between behavioral and ERP data

To test the functional interpretation of the targeted ERP effects, i.e., the greater availability of cognitive resources during higher P3 amplitudes, we used LME regression models. P3 mean amplitudes were regressed on the following variables: RTs, the experimental variable (Trial type for switch costs and Block type for mixing costs), and their interaction. All models included Participants as crossed random effects and the experimental variable as a slope. The fitting procedure and contrasts were as the ones described for the switch and mixing costs models in section 5.3.3.3.4. We expected that RTs would show a significant interaction with the experimental variable, so faster reaction times would be related to larger ERP effects. We used two models, one for switch costs and one for mixing costs. The mixing cost model only converged when the Block type slope was removed from the Participant intercept. The switch cost model showed only a significant effect for RTs (t = -4.56). The mixing model showed the main effects of RTs (t = -8.28), Block type (t = 4.05), and the interaction between RTs and Block type (t = 3.88). In both models, the significant effect of RTs was expressed as an inverse relationship between P3 mean amplitude and RTs: shorter reaction times predicted a more positive P3 mean amplitude. The significant interaction between RTs and Block type in the mixing cost model was the result of larger mixing costs for shorter reaction times compared to longer reaction times.

5.5 Discussion

The current study aimed to investigate the effect of external constraints on language switching on the transfer of domain-general cognitive abilities. We tested bilingual participants in a training experiment, where half completed a forced language switching task and the other half a voluntary one. Before and after the training, the participants performed a non-linguistic task switching task. We predicted a reduction in switch costs during the post-test of the non-linguistic task switching for the forced language switching group compared to the voluntary switching group and the pre-test test phase. For mixing costs, we expected no group differences between test phases and groups. We found that participants from both groups improved both their switch and mixing costs from pre- to post-test. However, the change in switch costs was not modulated by group. Instead, the groups differed in their mixing cost changes, with the forced language switching group showing a greater improvement in mixing costs than the voluntary language switching group. These behavioral results were supported by similar results for the P3 mean amplitude, with the voluntary language switching group showing greater mixing costs during the post-test than the forced language switching group. Together these results indicate that forced language switching engages more global cognitive control mechanisms than voluntary language switching.

5.5.1 Switch costs

Consistent with our expected results, we found an effect of training in the behavioral data for switch costs, demonstrating that the participants were generally faster during the post-test and benefitted from a task repetition effect. Additionally, there was a typical switch cost with reaction times being longer for switch trials than repeat trials (Karayanidis et al., 2003, 2011; Kieffaber & Hetrick, 2005; Lorist et al., 2000; Nicholson et al., 2005; Poulsen et al., 2005). While we initially predicted reduced switch costs for the forced language switching group in the post-test, we did not find evidence for transient trial-to-trial cognitive processes as switch costs showed no transfer from linguistic training to non-linguistic task switching as predicted by the ACH (Green & Abutalebi, 2013). Our findings suggest that both forced and voluntary language switching engage similar mechanisms for actively and rapidly switching between two languages. This lack of differentiation in transient control mechanisms implies that the cognitive processes involved in language switching, such as response inhibition and task engagement and disengagement, are comparable regardless of whether language choice is limited by external constraints or not. The results demonstrate that language switches, regardless of how they were initiated, require transient cognitive control regardless of how they were initiated. Even when the voluntary language switching group showed smaller switch costs in the language switching task, their transfer of transient cognitive control was similar to the forced language switching group. The apparent reduced difficulty in voluntary language switching might therefore not be the result of reduced cognitive control, compared to forced language switching, but instead caused by a greater reliance on bottom-up processes like lexical frequency.

Similar to the behavioral data, our P3 mean amplitudes only responded to the repetition effect and switch costs, demonstrating the validity of our task design and giving no evidence for transfer of cognitive mechanisms from language switching to non-linguistic task switching. Interestingly, the P3 mean amplitude was more positive for switch trials than for repeat trials. The direction of P3 amplitude in our switch costs aligns with previous studies on non-linguistic task switching with a more positive P3 amplitude for switch trials than repeat trials (Barcelo et al., 2006; Periáñez & Barceló, 2009; Timmer et al., 2017). As our design was closely modeled after Timmer et al. (2017), this similarity was to be expected and confirms the validity of the design. However, other studies have found diverging results for the P3 amplitude direction with more positive amplitudes for repeat trials than switch trials (Gajewski et al., 2010; Goffaux et al., 2006; Karayanidis et al., 2011; López Zunini et al., 2019; Nicholson et al., 2005; Petruo et al., 2019; Richardson et al., 2022; Whitson et al., 2014; Wong et al., 2018). P3 has been shown to be sensitive to the specific details of the experimental design. The cue stimulus interval (Nicholson et al., 2005; Wylie et al., 2009), frequency of response types (Zhuo et al., 2021), and the response modality (Hsieh et al., 2014) have all been shown to influence the amplitude of the P3 component. Furthermore, both the presentation of the cue as well as the stimulus presentation elicit P3-like ERP components (Jost et al., 2008; Kieffaber & Hetrick, 2005) that can overlap with small cue stimulus intervals. The current study presents cue and stimulus simultaneously, which might cause an overlap of functionally different ERP components and affect the results. Interestingly, Kieffaber and Hetrick (2005) found that with a cue stimulus interval of 1200 ms, more positive amplitudes for switch trials for a cue-bound P3 component, while their target-bound P3 had more positive amplitudes for repeat trials. We would therefore suggest that the direction of the P3 mean amplitude can be attributed to the unique features of our experimental design, reproduced from Timmer et al. (2017), and the overlap of functionally different ERP components. This overlap of contradictory components would potentially reduce the visibility of switch costs from the electrophysiological signal and obscure any potential group differences. Additionally, while we did find a significant relationship between behavioral data and P3 amplitudes, this link did not extend to switch costs as there was no significant interaction between RTs and Trial type for predicting the P3 amplitude (Section 5.4.3). Still, this lack of significant results in predicting P3 amplitude based on reaction times, might be the result of overlapping processes influencing the P3 amplitude, as described above.

5.5.2 Mixing costs

When comparing the pure and repeat trials, we found an effect of training, with an overall reduction of reaction times from pre-test to post-test. We also found typical mixing costs in the behavioral data, with longer reaction times for repeat trials than pure trials. The participants of both groups were also able to benefit from the repetition of the task and further reduce their mixing costs in the post-test. Critically, we also found smaller behavioral mixing costs during the post-test for the forced language switching group compared to the voluntary language switching group. These behavioral results coincide with the electrophysiological data, where we also found a group difference in the change of mixing costs between pre- and post-test. The voluntary language switching group showed greater mixing costs during the post-test compared to the forced language switching group. In contrast to our initial prediction, we found an effect of training group on mixing costs. This result provides evidence for transfer of global control mechanisms from linguistic training to a non-linguistic task, but only when language choice is restricted by language cues. This improvement in language mixing as the result of language training coincides with previous studies that suggested an increase in global control, such as conflict monitoring due to forced language switching training (Prior & Golan 2013, Liu et al. 2019). As mixing costs during non-linguistic task switching are generally assumed to reflect global monitoring processes (Braver et al., 2003; Prior & Macwhinney, 2010), our results reflect greater efficiency in managing multiple active task sets.

The direction of our P3 mixing costs aligns with previous studies, including Timmer et al. (2017). A common interpretation of this effect is that more positive P3 amplitudes in pure trials indicate the availability of more cognitive resources within the working memory for those trials (Goffaux et al., 2006; Kok, 2001). Our results show greater mixing costs for the voluntary language switching group during the post-test compared to the forced language switching group. This indicates that forced language switching training helps with the global managing of cognitive resources when task switching. While the voluntary language group had more cognitive resources available for the pure trials, their participants also performed worse in repeat trials (indicated by larger mixing costs in the P3 mean amplitude). The forced language switching group, on the other hand, was able to apply shared mechanisms from their language switching training and better allocate cognitive resources during task switching by improving their performance in repeat trials in the mixed block but also slightly decreasing it in pure trials (indicated by smaller mixing costs in the P3 mean amplitude). Overall, the forced language switching group managed their task switching task sets better and balanced cognitive resources between pure and repeat trials as a result of their language switching training.

5.5.3 Language control and domain-general cognitive control

The current study found a reduction in mixing costs after forced language switching training. Together with no effect on switch costs, this is evidence of an adaptation of global cognitive control but not of local cognitive control. That is, no short-term control processes are involved during the immediate presentation of external constraints on language choice adapted. However, the global management of multiple languages within the language system was activated differently during forced language switching compared to voluntary language switching. These results coincide with previous studies that found an advantage for bilinguals in conflict monitoring (Bialystok et al., 2005; Costa, Hernández, et al., 2009; Prior & Gollan, 2013). However, our results suggest that this is not the result of only bilingualism, i.e., the simultaneous activation of two languages simultaneously, but instead the result of external demands on the language system. While languages are regularly activated simultaneously and can cause interference with each other (Thierry & Wu, 2007), the involved cognitive control processes do not necessarily transfer to other non-linguistic domains. Instead, broader processes such as conflict monitoring (Costa et al., 2009; Soveri et al., 2011) are considered to be involved. Our results give insight into the possible link between language control and domain-general cognitive control. While previous studies found advantages for bilinguals in both *switching* and *mixing* over monolingual speakers (Bialystok et al., 2005; Costa, Hernández, et al., 2009; Soveri et al., 2011), these results were not always found (Lehtonen et al., 2018; Prior & Macwhinney, 2010). The present study suggests that language activation is not responsible for how bilinguals engage specific control abilities. While the transfer of efficiency in shared mechanisms between language switching and non-linguistic task switching is well known and thought to rely on shared mechanisms of alertness and conflict monitoring (Liu et al., 2019; Timmer et al., 2019), our results demonstrate that just using two languages interchangeably is not sufficient. Instead, bilinguals follow the specific constraints of an interactional context and adapt their cognitive system accordingly (Beatty-Martínez & Titone, 2021). The present data suggest that language switching under external constraints on language choice requires an increased need to maintain and regulate both languages. On the other hand, during voluntary language switching, this need for cognitive control is reduced, and bilinguals are free to follow their bottom-up processes, such as lexical access (de Bruin et al., 2018).

The current study found no support for the transfer of transient cognitive mechanisms like task engagement and disengagement predicted by the adaptive control hypothesis (Green & Abutalebi, 2013). Instead, we found that external constraints on language choice in language switching promote the use of global mechanisms like conflict monitoring. Even short-term training in different language contexts, in the form of language switching training, leads to an adaptation in the cognitive system to use mental resources more efficiently. So, while there was no transfer in the predicted cognitive mechanisms, we still found a close relationship between language control and domain-general cognitive control. This modulation of cognitive control by the language environment aligns with the general idea of the ACH. These findings also support the notion of Blanco-Elorrieta and Pylkkänen (2018) that not all forms of bilingual language switching lead to an advantage in domain-general cognitive control. Instead, only cognitive mechanisms shared between externally constrained language switching and non-linguistic task switching benefit from this transfer. Specifically, global processes such as conflict monitoring seem to be affected. Altogether, this study forms another set of evidence in the growing trend in bilingualism research not to treat bilinguals and their experiences as monolithic classes but instead carefully distinguish their specific interactional language contexts and how they can affect cognitive control (Beatty-Martínez et al., 2020; Castro et al., 2022; Kałamała et al., 2022; Luk & Esposito, 2020).

5.5.4 Limitations

One difficulty of our task design is the creation of a comparable language switching training that allows us to differentiate between the effects of language switching with or without external constraints on language choice. Because participants are completely free to switch languages within the voluntary language switching task, it becomes difficult to control their language use of L1 and L2 as well as their portion of switch trials within the task. In order to mitigate these problems, we carefully selected our stimuli to encourage equal language switches for all pictures and remove any potential confounds, such as lexical frequency or name agreement. Additionally, we measured the participants' heart rates during the task and asked them to rate their perceived effort to assess the difference in difficulty between the forced and voluntary language switching tasks.

We found both average proportions of English trials and proportions of switch trials within the voluntary language switching group comparable to the forced language switching group, indicating a similar behavior in language use for both language switching groups. However, the voluntary language switching group demonstrated a larger variability among participants. The switching frequency of 38% observed during our experiment is comparable to previous studies investigating voluntary language switching (Blanco-Elorrieta & Pylkkänen, 2017; de Bruin et al., 2018; Gollan et al., 2014; Kleinman & Gollan, 2016). For our additional measures, we found no difference in heart rate between the two types of training nor an effect on perceived effort. Given the differences in the number of switch trials between forced and voluntary language switching, the argument could be made that our voluntary language switching task is somewhere between our forced language switching task and a block of single language picture naming, at least for some participants. Our results would then coincide with previous studies (Liu et al., 2019; Timmer et al., 2019), especially as they, too, found effects of training on non-linguistic task mixing. However, given that the number of English trials is very similar between the forced and voluntary language switching groups, we would argue that we still are able to differentiate between the effects of language choice and the overall presence of two languages. Given the similar amount of L1 and L2 use within a short period of time for both groups, the participants would show similar levels of language activation and engage similar cognitive mechanisms of language control to manage this co-activation. However, the greater use of global cognitive control by the forced language switching group cannot be attributed to this type of language control and was instead enhanced by the presence of external demands on the language system in the form of language cues. When considering both our results in the transfer of cognitive control efficiency and the similar behavior of both groups in regard to language switching, we would argue that the external constraints within the forced language switching training are the most likely factor responsible for the group differences and not differences in switching rate or percentage of L2 use.

5.6 Conclusion

The present study revealed that the presence of constraints on language choice in an interactional context modulates the use of global cognitive control. The effects of those constraints can be observed after short-term language switching training in an experimental setting as efficiency in conflict monitoring was transferred from the linguistic to the non-linguistic domain. These findings show the ability of the general cognitive system to adapt rapidly to subtle differences in the language environment and how the difficulty of using two languages can differ as a result.

6 General Discussion

Research on the cognitive consequences of bilingualism has produced a wide variety of results, with some seeing a positive relationship between bilingualism and cognitive control (Bialystok, 2009, 2017; Costa et al., 2008), while others could find no such causal link (Hernández et al., 2013; Paap et al., 2015; Paap & Greenberg, 2013). While a definite answer to whether the bilingual advantage exists still eludes us, researchers have started to take a more fine-grained look at the bilingual experience itself and tried to establish what factors in someone's language environment would promote the use of cognitive control in different ways (Bak, 2016; Beatty-Martínez & Titone, 2021; D. W. Green & Abutalebi, 2013; D. W. Green & Wei, 2014; Gullifer & Titone, 2020; Kałamała et al., 2023; Navarro-Torres et al., 2021; Wodniecka, Casado, et al., 2020). This dissertation contains a series of investigations and analyses studying the effects of the language environment on language use and domain-general cognitive control. It attempted to probe different aspects of the interaction between environment and language control. Well-known paradigms such as picture naming and language switching were employed, and relatively simple processes like lexical access during word production were tested, which enabled a direct comparison of our outcomes with results from the previous literature. Instead of the traditional approach taken in many previous studies (i.e., comparing the performance of monolinguals and bilinguals), we tested only bilingual populations but varied and/or manipulated their language environment. In Investigations 1 and 2, we explored language control mechanisms in one group of bilinguals but in two different language contexts, i.e., during L2 immersion and soon after a brief L1 reimmersion. In Investigation 3, we used an experimentally induced language environment with a forced vs. free language switching task and measured the consequences of this manipulation on the efficiency of cognitive control. By carefully selecting and matching participants as closely as possible and combining within and between-group designs, it was possible to control many confounding factors, such as differences in language proficiency, socioeconomic status, or language switching behavior. These factorial designs revealed that even short-term environmental changes greatly influenced the participants' language use and cognitive control.

All three investigations revealed quick adaptations in the language system and cognitive control when aspects of the language environment change. In Investigations 1 and 2, we found that lexical L1 frequency significantly interacts with the changes in lexical access, most likely due to the different frequencies of occurrence of single words in different language environments and how bilinguals seem to apply proactive language control to combat L2 interference. Investigation 3 found that the restriction of language choice by the inclusion of language cues in a language switching paradigm significantly changes the use of global cognitive control, even when the participants showed similar language use with respect to the frequency of language switches and the amount of L1 and L2 words they produced. Together, these studies indicate that even short-term changes in the environment can influence how language control and cognitive control are engaged. The current results can be understood in the context of models of language use that assume that bilingualism is a complex network of experiences and dynamic changes with differences in language use and engagement of cognitive control. Hopefully, it would allow the identification of underlying processes and a better understanding of how the human language system operates and adapts rapidly. Below, I place the three investigations within the larger field of cognitive psychology and how they contribute to a better understanding of how to investigate the cognitive consequences of bilingualism.

6.1 Lexical access during L1 and L2 immersion

Investigations 1 and 2 focused on how changes in the language environment can influence lexical L1 access. Specifically, how long-term L2 immersion manifests and how malleable its effects might be. By combining both within- and between-group comparisons, we were able to eliminate confounding factors such as individual differences in language experience between groups. While Investigation 1 focuses on analyzing the behavioral data, Investigation 2 investigates the electrophysiological markers of lexical L1 access in the form of event-related potentials. Together the results of these investigations made it possible to study the different aspects of lexical access and get a more comprehensive understanding of how native word production operates under L1 or L2 immersion.

Investigation 1 demonstrated the capability of the language system to adapt to the constant L2 interferences of an L2 dominant environment. L2 immersed participants were able to maintain a level of lexical L1 access similar to participants under L1 immersion. However, when members of a migrant population return to their native language environment, the neurocognitive system undergoes quick adaptations; even by only short-term exposure to L1, migrants were able to reduce their cognitive resources allocated to language control, visible in the amplitude reduction of the P2 component. These results, therefore, form another piece of evidence for lexical L1 access during L2 immersion (Beatty-Martínez et al., 2020; Yilmaz & Schmid, 2012) and the flexibility of the overall language system (Baus et al., 2013; Linck et al., 2009). Investigation 2 gave us a potential explanation of the underlying mechanisms of the observed behavioral results. While we tried to examine established markers of the ease of lexical L1 access, we found a diverging ERP pattern instead. We interpreted these novel ERP results as additional language control employed by the migrant population during L2 immersion. As previous studies already found behavioral indicators of increased proactive control during L2 immersion, that is, greater use of goal maintenance or conflict monitoring (Beatty-Martínez et al., 2020; Zhang et al., 2021), we argued that our participants used similar cognitive mechanisms to combat L2 interference during L2 immersion. This increased use of proactive language control might explain how the long-term migrant population could maintain native-like lexical access within an L2-dominant environment. When returning to an L1-dominant environment, this proactive language control was presumably replaced by a different type of language control to benefit from the higher occurrence of L1 cues. More reactive forms of language control allow quicker responses to language cues from the environment (Beatty-Martínez et al., 2020), especially to those words with a higher frequency of use, as indicated by our observed significant effect of items with high lexical frequency. However, future studies are needed to confirm these hypotheses and fully understand the involved processes.

6.2 Cognitive control and external constraints

In Investigation 3, we focused on the domain-general cognitive abilities of bilinguals. Following suggestions to resolve the discussion surrounding the bilingual advantage (Beatty-Martínez & Titone, 2021; Blanco-Elorrieta & Pylkkänen, 2018; D. W. Green & Abutalebi, 2013), we compared the impact of environmental constraints on language switching on cognitive control. We developed a novel design that combined within- and between-group comparisons in a training study using cued and voluntary language switching. By examining the transfer of cognitive efficiency from language switching to non-linguistic task switching, we demonstrated that external constraints on language choice encourage greater use of global control processes. That is, the use of language cues in language switching resulted in smaller mixing costs in a successive non-linguistic task switching test. Although previous studies already found that voluntary language switching requires only reduced cognitive control (Blanco-Elorrieta & Pylkkänen, 2017; Hartanto & Yang, 2016), the presented research is one of the first to demonstrate this effect in both behavioral and electrophysiological data using a factorial design. In particular, the results of Investigation 3 should inform the development of future experiments investigating the effects of bilingualism. When conducting an experimental study involving language switching, one should be careful which factors in a given study contribute to the observed effects. As Blanco-Elorrieta and Pylkkänen (2018) argue in their review, some of the variability in the research of the bilingual advantage might not be the result of simply bilingualism, fluency in two or more languages but instead caused by artificial demands within the experimental task. They argue that future studies should involve more ecological designs, specifically, language switches not controlled by simple language cues but instead by natural markers like human faces (Blanco-Elorrieta & Pylkkänen, 2017) or conversational partners (Kałamała et al., 2021). The results of Investigation 3 demonstrate the need for this type of new experimental design and how artificial language environments could induce short-term changes in cognitive control. On a broader scale, Investigation 3 demonstrates the relevance of adequately characterizing the language switching experience of bilinguals when exploring their cognitive control. Many questionnaires and measures have recently been developed to accommodate this need for quantifying switching behavior (Gullifer & Titone, 2020; Kałamała et al., 2023). By accounting for participants' individual language switching experiences, future studies might be able to develop more precise techniques and designs for studying the bilingual advantage.

6.3 Limitations

One problem of the present studies is the task impurity problem, the idea that all cognitive tasks inevitably involve complex sets of cognitive mechanisms rather than measuring a single multipurpose mechanism (Friedman, 2016; Kałamała et al., 2020; Miyake et al., 2000). While mixing costs are thought to reflect more proactive and sustained mechanisms like cue monitoring (Braver et al., 2007; Kiesel et al., 2010; Rubin & Meiran, 2005), they are not necessarily related to the exact underlying cognitive mechanisms as the proactive control measured with the AX-CPT task, given that many studies on cognitive control often found low correlations between cognitive tasks that are supposed to tap into the same cognitive mechanisms (Miyake et al., 2000), and specifically low correlations between task switching and the AX-CPT (Snijder et al., 2023). The low correlation between measures of proactive language control and mixing costs in task switching complicate establishing shared cognitive mechanisms between tasks and experiments. Task impurity has been argued before to contribute to the variation in study results of the bilingual advantage (Friedman, 2016; Paap et al., 2015; Valian, 2015), so even with detailed knowledge of the language experience of participants and a working theory of language contexts, it would be challenging to isolate the affected cognitive mechanisms, if there are no consistent measures across tasks.

6.4 Conclusion and perspectives

The presented studies contribute to the field of cognitive psychology in several aspects. Investigations 1 and 2 gave us a better understanding of how migrants could expect to maintain the availability of their native language when moving to a different country for the foreseeable future. While their language does seem to apply additional language control to combat L2 interference, this does not manifest in overtly adverse outcomes for their language use. Instead, they will be able to maintain their native language use and quickly adapt to their surroundings when visiting their native language environments. This recovery is possible thanks to the adaptive capabilities of the language system and its ability to respond rapidly to any environment.

Investigation 3 expanded on this adaption to the environments by focusing on specific external constraints on the language system. Using the pre-posttest design, we demonstrated that external constraints encourage greater use of global cognitive control processes as bilinguals have a greater need for cue monitoring in order to conform to external demands. Investigations 1, 2, and 3 give a picture of a rapidly adapting language system that changes the use of cognitive control based on even minor environmental changes. Through this, the dissertation supports treating bilingualism as a collection of different language contexts and experiences. By considering these ideas, it should be possible to establish more specialized experimental tasks to examine the effects of language experience on cognitive control in the future.

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.

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140–155. https://doi.org/10.3758/BF03195559

Ammerlaan, T. (1996). *“You get a bit wobbly...”: Exploring bilingual lexical retrieval processes in the context of first language attrition*. [Sl: sn].

Bae, S., & Masaki, H. (2019). Effects of Acute Aerobic Exercise on Cognitive Flexibility Required During Task-Switching Paradigm. *Frontiers in Human Neuroscience*, *13*. https://www.frontiersin.org/articles/10.3389/fnhum.2019.00260

Bak, T. H. (2016). Cooking pasta in La Paz. *Linguistic Approaches to Bilingualism*, *6*(5), 699–717.

Barcelo, F., Escera, C., Corral, M. J., & Periáñez, J. A. (2006). Task Switching and Novelty Processing Activate a Common Neural Network for Cognitive Control. *Journal of Cognitive Neuroscience*, *18*(10), 1734–1748. https://doi.org/10.1162/jocn.2006.18.10.1734

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

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, *14*(2), 201–212. https://doi.org/10.1016/0278-2626(90)90029-N

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious Mixed Models. *ArXiv:1506.04967 [Stat]*. http://arxiv.org/abs/1506.04967

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

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., Herron, D., Ching Lu, C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., … Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*(2), 344–380. https://doi.org/10.3758/BF03196494

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

Baus, C., Santesteban, M., Runnqvist, E., Strijkers, K., & Costa, A. (2020). Characterizing lexicalization and self-monitoring processes in bilingual speech production. *Journal of Neurolinguistics*, *56*, 100934. https://doi.org/10.1016/j.jneuroling.2020.100934

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), Article 4. https://doi.org/10.3390/languages6040168

Bergmann, C., Meulman, N., Stowe, L. A., Sprenger, S. A., & Schmid, M. S. (2015). Prolonged L2 immersion engenders little change in morphosyntactic processing of bilingual natives. *Neuroreport*, *26*(17), 1065–1070.

Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism*, *10*(3), 210–223.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent*. *Bilingualism: Language and Cognition*, *12*(1), 3–11. https://doi.org/10.1017/S1366728908003477

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, *143*(3), 233–262. https://doi.org/10.1037/bul0000099

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism*, *9*(1), 103–119. https://doi.org/10.1177/13670069050090010701

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

Blanco-Elorrieta, E., & Pylkkänen, L. (2017). Bilingual Language Switching in the Laboratory versus in the Wild: The Spatiotemporal Dynamics of Adaptive Language Control. *Journal of Neuroscience*, *37*(37), 9022–9036. https://doi.org/10.1523/JNEUROSCI.0553-17.2017

Blanco-Elorrieta, E., & Pylkkänen, L. (2018). Ecological Validity in Bilingualism Research and the Bilingual Advantage. *Trends in Cognitive Sciences*, *22*(12), 1117–1126. https://doi.org/10.1016/j.tics.2018.10.001

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 (Cambridge, England)*, *22*(5), 1052–1067. https://doi.org/10.1017/S1366728918000858

Botezatu, M. R., Kroll, J. F., Trachsel, M. I., & Guo, T. (2020). *Second language immersion impacts native language lexical production and comprehension*. https://doi.org/10.1075/lab.19059.bot

Branzi, F. M., Martin, C. D., Abutalebi, J., & Costa, A. (2014). The after-effects of bilingual language production. *Neuropsychologia*, *52*, 102–116. https://doi.org/10.1016/j.neuropsychologia.2013.09.022

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., Janowsky, J. S., Taylor, S. F., Yesavage, J. A., Mumenthaler, M. S., Jagust, W. J., & Reed, B. R. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, *130*(4), 746–763. https://doi.org/10.1037/0096-3445.130.4.746

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.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron*, *39*(4), 713–726.

Calabria, M., Hernandez, M., Branzi, F., & Costa, A. (2012). Qualitative Differences between Bilingual Language Control and Executive Control: Evidence from Task-Switching. *Frontiers in Psychology*, *2*. https://www.frontiersin.org/articles/10.3389/fpsyg.2011.00399

Casado, A., Walther, J., Wolna, A., Szewczyk, J., Sorace, A., & Wodniecka, Z. (2023). Advantages of visiting your home country: How brief reimmersion in their native country impacts migrants' native language access. *Bilingualism: Language and Cognition*, 1–12. https://doi.org/10.1017/S136672892300024X

Castro, S., Wodniecka, Z., & Timmer, K. (2022). Am I truly monolingual? Exploring foreign language experiences in monolinguals. *PLOS ONE*, *17*(3), e0265563. https://doi.org/10.1371/journal.pone.0265563

Chamorro, G., & Sorace, A. (2019). The Interface Hypothesis as a framework for studying L1 attrition. In B. Köpke & M. Schmid (Eds.), *The Oxford Handbook of Language Attrition*. Oxford University Press. https://kar.kent.ac.uk/66124/

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

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, *1147*, 192–208. https://doi.org/10.1016/j.brainres.2007.01.137

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135–149. https://doi.org/10.1016/j.cognition.2009.08.001

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59–86. https://doi.org/10.1016/j.cognition.2006.12.013

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. https://doi.org/10.1073/pnas.0908921106

Cuetos, F., Ellis, A. W., & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers*, *31*(4), 650–658. https://doi.org/10.3758/BF03200741

Curran, T., Tanaka, J. W., & Weiskopf, D. M. (2002). An electrophysiological comparison of visual categorization and recognition memory. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(1), 1–18. https://doi.org/10.3758/CABN.2.1.1

de Bruin, A., Samuel, A. G., & Duñabeitia, J. A. (2018). Voluntary language switching: When and why do bilinguals switch between their languages? *Journal of Memory and Language*, *103*, 28–43. https://doi.org/10.1016/j.jml.2018.07.005

Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review*, *22*(6), 1630–1645. https://doi.org/10.3758/s13423-015-0836-1

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

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. https://doi.org/10.1016/j.neuroimage.2006.11.004

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

Eddy, M., Schmid, A., & Holcomb, P. J. (2006). Masked repetition priming and event-related brain potentials: A new approach for tracking the time-course of object perception. *Psychophysiology*, *43*(6), 564–568. https://doi.org/10.1111/j.1469-8986.2006.00455.x

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 202–224. https://doi.org/10.1037/0278-7393.27.1.202

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

Friedman, N. P. (2016). Research on individual differences in executive functions: Implications for the bilingual advantage hypothesis. *Linguistic Approaches to Bilingualism*, *6*(5), 535–548. https://doi.org/10.1075/lab.15041.fri

Gajewski, P. D., & Falkenstein, M. (2011). Diversity of the P3 in the task-switching paradigm. *Brain Research*, *1411*, 87–97. https://doi.org/10.1016/j.brainres.2011.07.010

Gajewski, P. D., Kleinsorge, T., & Falkenstein, M. (2010). Electrophysiological correlates of residual switch costs. *Cortex*, *46*(9), 1138–1148. https://doi.org/10.1016/j.cortex.2009.07.014

Gajewski, P., & Falkenstein, M. (2012). Training-Induced Improvement of Response Selection and Error Detection in Aging Assessed by Task Switching: Effects of Cognitive, Physical, and Relaxation Training. *Frontiers in Human Neuroscience*, *6*. https://www.frontiersin.org/articles/10.3389/fnhum.2012.00130

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

Gehring, W. J., Bryck, R. L., Jonides, J., Albin, R. L., & Badre, D. (2003). The mind’s eye, looking inward? In search of executive control in internal attention shifting. *Psychophysiology*, *40*(4), 572–585. https://doi.org/10.1111/1469-8986.00059

Gnitiev, S., & Bátyi, S. (2022). Lexical access, lexical diversity and speech fluency in first language attrition. *Strani Jezici: Časopis Za Primijenjenu Lingvistiku*, *51*(2), 159–184.

Goffaux, P., Phillips, N. A., Sinai, M., & Pushkar, D. (2006). Behavioural and electrophysiological measures of task switching during single and mixed-task conditions. *Biological Psychology*, *72*(3), 278–290. https://doi.org/10.1016/j.biopsycho.2005.11.009

Gollan, T. H., & Ferreira, V. S. (2009). Should I Stay or Should I Switch? A Cost–Benefit Analysis of Voluntary Language Switching in Young and Aging Bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(3), 640–665. https://doi.org/10.1037/a0014981

Gollan, T. H., Kleinman, D., & Wierenga, C. E. (2014). What's Easier: Doing What You Want, or Being Told What to Do? Cued versus Voluntary Language and Task Switching. *Journal of Experimental Psychology. General*, *143*(6), 2167–2195. https://doi.org/10.1037/a0038006

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

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186.

Gratton, C., Evans, K. M., & Federmeier, K. D. (2009). See what I mean? An ERP study of the effect of background knowledge on novel object processing. *Memory & Cognition*, *37*(3), 277–291. https://doi.org/10.3758/MC.37.3.277

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

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Green, D. W., & Wei, L. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience*, *29*(4), 499–511. https://doi.org/10.1080/23273798.2014.882515

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Gross, M., & Kaushanskaya, M. (2015). Voluntary language switching in English–Spanish bilingual children. *Journal of Cognitive Psychology*, *27*(8), 992–1013. https://doi.org/10.1080/20445911.2015.1074242

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, *23*(2), 283–294. https://doi.org/10.1017/S1366728919000026

Haman, E., Łuniewska, M., Hansen, P., Simonsen, H. G., Chiat, S., Bjekić, J., Blažienė, A., Chyl, K., Dabašinskienė, I., Abreu, P. E. de, Gagarina, N., Gavarró, A., Håkansson, G., Harel, E., Holm, E., Kapalková, S., Kunnari, S., Levorato, C., Lindgren, 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

Hartanto, A., & Yang, H. (2016). Disparate bilingual experiences modulate task-switching advantages: A diffusion-model analysis of the effects of interactional context on switch costs. *Cognition*, *150*, 10–19. https://doi.org/10.1016/j.cognition.2016.01.016

Henson, R. N., Rylands, A., Ross, E., Vuilleumeir, P., & Rugg, M. D. (2004). The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *NeuroImage*, *21*(4), 1674–1689. https://doi.org/10.1016/j.neuroimage.2003.12.020

Hernández, M., Martin, C. D., Barceló, F., & Costa, A. (2013). Where is the bilingual advantage in task-switching? *Journal of Memory and Language*, *69*(3), 257–276. https://doi.org/10.1016/j.jml.2013.06.004

Holcomb, P. J., & Mcpherson, W. B. (1994). Event-Related Brain Potentials Reflect Semantic Priming in an Object Decision Task. *Brain and Cognition*, *24*(2), 259–276. https://doi.org/10.1006/brcg.1994.1014

Hopp, H., & Schmid, M. S. (2013). Perceived foreign accent in first language attrition and second language acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics*, *34*(2), 361–394. https://doi.org/10.1017/S0142716411000737

Hsieh, S., & Liu, H. (2009). Electrophysiological evidence of the adaptive task-set inhibition in task switching. *Brain Research*, *1255*, 122–131. https://doi.org/10.1016/j.brainres.2008.11.103

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

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

Inquisit 5 [Computer software]. (2016). *Retrieved from https://www.millisecond.com*.

Jacobs, A., Fricke, M., & Kroll, J. F. (2016). Cross-Language Activation Begins During Speech Planning and Extends Into Second Language Speech. *Language Learning*, *66*(2), 324–353. https://doi.org/10.1111/lang.12148

Jevtović, M., Duñabeitia, J. A., & Bruin, A. de. (2020). How do bilinguals switch between languages in different interactional contexts? A comparison between voluntary and mandatory language switching. *Bilingualism: Language and Cognition*, *23*(2), 401–413. https://doi.org/10.1017/S1366728919000191

Jost, K., Mayr, U., & Rösler, F. (2008). Is task switching nothing but cue priming? Evidence from ERPs. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(1), 74.

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. https://doi.org/10.1016/S1388-2457(00)00386-2

Kałamała, P., Chuderski, A., Szewczyk, J., Senderecka, M., & Wodniecka, Z. (2023). Bilingualism caught in a net: A new approach to understanding the complexity of bilingual experience. *Journal of Experimental Psychology: General*, *152*(1), 157–174. https://doi.org/10.1037/xge0001263

Kałamała, P., Senderecka, M., & Wodniecka, Z. (2022). On the multidimensionality of bilingualism and the unique role of language use. *Bilingualism: Language and Cognition*, *25*(3), 471–483. https://doi.org/10.1017/S1366728921001073

Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition*, *204*, 104373. https://doi.org/10.1016/j.cognition.2020.104373

Kałamała, P., Walther, J., Zhang, H., Diaz, M., Senderecka, M., & Wodniecka, Z. (2021). The use of a second language enhances the neural efficiency of inhibitory control: An ERP study. *Bilingualism: Language and Cognition*, 1–18. https://doi.org/10.1017/S1366728921000389

Kamijo, K., & Takeda, Y. (2010). Regular physical activity improves executive function during task switching in young adults. *International Journal of Psychophysiology*, *75*(3), 304–311. https://doi.org/10.1016/j.ijpsycho.2010.01.002

Karayanidis, F., Coltheart, M., Michie, P. T., & Murphy, K. (2003). Electrophysiological correlates of anticipatory and poststimulus components of task switching. *Psychophysiology*, *40*(3), 329–348. https://doi.org/10.1111/1469-8986.00037

Karayanidis, F., Whitson, L., Heathcote, A., & Michie, P. (2011). Variability in Proactive and Reactive Cognitive Control Processes Across the Adult Lifespan. *Frontiers in Psychology*, *2*. https://www.frontiersin.org/articles/10.3389/fpsyg.2011.00318

Kieffaber, P. D., & Hetrick, W. P. (2005). Event-related potential correlates of task switching and switch costs. *Psychophysiology*, *42*(1), 56–71. https://doi.org/10.1111/j.1469-8986.2005.00262.x

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin*, *136*, 849–874. https://doi.org/10.1037/a0019842

Kleinman, D., & Gollan, T. H. (2016). Speaking Two Languages for the Price of One: Bypassing Language Control Mechanisms via Accessibility-Driven Switches. *Psychological Science*, *27*(5), 700–714. https://doi.org/10.1177/0956797616634633

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, *38*(3), 557–577. https://doi.org/10.1017/S0048577201990559

Köpke, B., & Genevska-Hanke, D. (2018). First Language Attrition and Dominance: Same 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).

Kroll, J. F., Bobb, S. C., & Hoshino, N. (2014). Two Languages in Mind: Bilingualism as a Tool to Investigate Language, Cognition, and the Brain. *Current Directions in Psychological Science*, *23*(3), 159–163. https://doi.org/10.1177/0963721414528511

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, *144*(4), 394–425. https://doi.org/10.1037/bul0000142

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. https://doi.org/10.3758/s13428-011-0146-0

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

Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*(3), 673–680. https://doi.org/10.1017/S1366728913000606

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

Liu, C., Yang, C.-L., Jiao, L., Schwieter, J. W., Sun, X., & Wang, R. (2019). Training in Language Switching Facilitates Bilinguals' Monitoring and Inhibitory Control. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01839

López Zunini, R. A., Morrison, C., Kousaie, S., & Taler, V. (2019). Task switching and bilingualism in young and older adults: A behavioral and electrophysiological investigation. *Neuropsychologia*, *133*, 107186. https://doi.org/10.1016/j.neuropsychologia.2019.107186

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. https://doi.org/10.3389/fnhum.2014.00213

Lorist, M. M., Klein, M., Nieuwenhuis, S., Jong, R. D., Mulder, G., & Meijman, T. F. (2000). Mental fatigue and task control: Planning and preparation. *Psychophysiology*, *37*(5), 614–625. https://doi.org/10.1111/1469-8986.3750614

Luk, G., & Esposito, A. G. (2020). BLC mini-series: Tools to document bilingual experiences. *Bilingualism: Language and Cognition*, *23*(5), 927–928. https://doi.org/10.1017/S1366728920000632

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

Marecka, M., Wrembel, M., Otwinowska, A., Szewczyk, J., Banasik-Jemielniak, N., & Wodniecka, Z. (2020). BILINGUAL CHILDREN'S PHONOLOGY SHOWS EVIDENCE OF TRANSFER, BUT NOT DECELERATION IN THEIR L1. *Studies in Second Language Acquisition*, *42*(1), 89–114. https://doi.org/10.1017/S0272263119000408

Marian Viorica, Blumenfeld Henrike K., & Kaushanskaya Margarita. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*(7), 703–704.

Meuter, R. F. I., & Allport, A. (1999). Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language*, *40*(1), 25–40. https://doi.org/10.1006/jmla.1998.2602

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *ArXiv:1301.3781 [Cs]*. http://arxiv.org/abs/1301.3781

Mitchell, D. B., & Brown, A. S. (1988). Persistent repetition priming in picture naming and its dissociation from recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 213–222. https://doi.org/10.1037/0278-7393.14.2.213

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Navarro-Torres, C. A., Beatty-Martínez, A. L., Kroll, J. F., & Green, D. W. (2021). Research on bilingualism as discovery science. *Brain and Language*, *222*, 105014. https://doi.org/10.1016/j.bandl.2021.105014

Nicholson, R., Karayanidis, F., Poboka, D., Heathcote, A., & Michie, P. T. (2005). Electrophysiological correlates of anticipatory task-switching processes. *Psychophysiology*, *42*(5), 540–554. https://doi.org/10.1111/j.1469-8986.2005.00350.x

Ooi, S. H., Goh, W. D., Sorace, A., & Bak, T. H. (2018). From Bilingualism to Bilingualisms: Bilingual experience in Edinburgh and Singapore affects attentional control differently. *Bilingualism: Language and Cognition*, *21*(4), 867–879. https://doi.org/10.1017/S1366728918000020

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(37), 15583–15587. https://doi.org/10.1073/pnas.0903620106

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*(2), 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014

Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., & Sawi, O. M. (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *Journal of Cognitive Psychology*, *29*(2), 89–112. https://doi.org/10.1080/20445911.2016.1248436

Periáñez, J. A., & Barceló, F. (2009). Updating sensory versus task representations during task-switching: Insights from cognitive brain potentials in humans. *Neuropsychologia*, *47*(4), 1160–1172. https://doi.org/10.1016/j.neuropsychologia.2009.01.014

Petruo, V. A., Mückschel, M., & Beste, C. (2019). Numbers in action during cognitive flexibility – A neurophysiological approach on numerical operations underlying task switching. *Cortex*, *120*, 101–115. https://doi.org/10.1016/j.cortex.2019.03.017

Philiastides, M. G., Ratcliff, R., & Sajda, P. (2006). Neural Representation of Task Difficulty and Decision Making during Perceptual Categorization: A Timing Diagram. *Journal of Neuroscience*, *26*(35), 8965–8975. https://doi.org/10.1523/JNEUROSCI.1655-06.2006

Philiastides, M. G., & Sajda, P. (2006). Causal Influences in the Human Brain During Face Discrimination: A Short-Window Directed Transfer Function Approach. *IEEE Transactions on Biomedical Engineering*, *53*(12), 2602–2605. https://doi.org/10.1109/TBME.2006.885122

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Poulsen, C., Luu, P., Davey, C., & Tucker, D. M. (2005). Dynamics of task sets: Evidence from dense-array event-related potentials. *Cognitive Brain Research*, *24*(1), 133–154. https://doi.org/10.1016/j.cogbrainres.2005.01.008

Prior, A., & Gollan, T. H. (2013). The elusive link between language control and executive control: A case of limited transfer. *Journal of Cognitive Psychology*, *25*(5), 622–645. https://doi.org/10.1080/20445911.2013.821993

Prior, A., & Macwhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, *13*(2), 253–262. https://doi.org/10.1017/S1366728909990526

R Development Core Team. (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.* www.R-project.org

Richardson, D. P., Foxe, J. J., Mazurek, K. A., Abraham, N., & Freedman, E. G. (2022). Neural markers of proactive and reactive cognitive control are altered during walking: A Mobile Brain-Body Imaging (MoBI) study. *NeuroImage*, *247*, 118853. https://doi.org/10.1016/j.neuroimage.2021.118853

Roux, F., Armstrong, B. C., & Carreiras, M. (2017). Chronset: An automated tool for detecting speech onset. *Behavior Research Methods*, *49*(5), 1864–1881. https://doi.org/10.3758/s13428-016-0830-1

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1477.

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. https://doi.org/10.1017/S1366728909990514

Schendan, H. E., & Kutas, M. (2002). Neurophysiological evidence for two processing times for visual object identification. *Neuropsychologia*, *40*(7), 931–945. https://doi.org/10.1016/S0028-3932(01)00176-2

Schendan, H. E., & Kutas, M. (2007). Neurophysiological Evidence for the Time Course of Activation of Global Shape, Part, and Local Contour Representations during Visual Object Categorization and Memory. *Journal of Cognitive Neuroscience*, *19*(5), 734–749. https://doi.org/10.1162/jocn.2007.19.5.734

Schmid, M. S. (2004). *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. John Benjamins Publishing.

Schmid, M. S. (2009). On L1 attrition and the linguistic system. *EUROSLA Yearbook*, *9*(1), 212–244. https://doi.org/10.1075/eurosla.9.11sch

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

Schneider, W. (1995). Micro experimental laboratory. *Pittsburg, PA: Psychology Software Tools*.

Snijder, J.-P., Tang, R., Bugg, J. M., Conway, A. R. A., & Braver, T. S. (2023). On the psychometric evaluation of cognitive control tasks: An Investigation with the Dual Mechanisms of Cognitive Control (DMCC) battery. *Behavior Research Methods*, 1–36. https://doi.org/10.3758/s13428-023-02111-7

Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–33. https://doi.org/10.1075/lab.1.1.01sor

Sorace, A. (2016). Referring expressions and executive functions in bilingualism. *Linguistic Approaches to Bilingualism*, *6*(5), 669–684. https://doi.org/10.1075/lab.15055.sor

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, *22*(3), 339–368. https://doi.org/10.1191/0267658306sr271oa

Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is There a Relationship between Language Switching and Executive Functions in Bilingualism? Introducing a within group Analysis Approach. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00183

Stolberg, D., & Münch, A. (2010). “Die Muttersprache vergisst man nicht” – or do you? A case study in L1 attrition and its (partial) reversal. *Bilingualism: Language and Cognition*, *13*(1), 19–31. https://doi.org/10.1017/S1366728909990332

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. https://doi.org/10.1016/j.bandl.2013.07.008

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. https://doi.org/10.1093/cercor/bhp153

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. https://doi.org/10.1016/j.jml.2011.06.002

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

Timmer, K., Calabria, M., & Costa, A. (2019). Non-linguistic effects of language switching training. *Cognition*, *182*, 14–24. https://doi.org/10.1016/j.cognition.2018.09.001

Timmer, K., Grundy, J. G., & Bialystok, E. (2017). Earlier and more distributed neural networks for bilinguals than monolinguals during switching. *Neuropsychologia*, *106*, 245–260. https://doi.org/10.1016/j.neuropsychologia.2017.09.017

Travers, S., & West, R. (2008). Neural correlates of cue retrieval, task set reconfiguration, and rule mapping in the explicit cue task switching paradigm. *Psychophysiology*, *45*(4), 588–601. https://doi.org/10.1111/j.1469-8986.2008.00658.x

Tsai, C., Pan, C., Chen, F., Wang, C., & Chou, F. (2016). The effects of acute aerobic exercise on a task-switching paradigm and BDNF levels in young adults with different levels of cardiorespiratory fitness. *Exp. Physiol*, *101*, 836–850.

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

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*(1), 3–24. https://doi.org/10.1017/S1366728914000522

Van Assche, E., Duyck, W., & Gollan, T. H. (2013). Whole-language and item-specific control in bilingual language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1781.

van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., Barisch, P., Vermeire, K., Lee, S.-H., & Lim, S. (2019). Does the Bilingual Advantage in Cognitive Control Exist and If So, What Are Its Modulating Factors? A Systematic Review. *Behavioral Sciences*, *9*(3), 27. https://doi.org/10.3390/bs9030027

van Petten, C., Senkfor, A. J., & Newberg, W. M. (2000). Memory for drawings in locations: Spatial source memory and event-related potentials. *Psychophysiology*, *37*(4), 551–564. https://doi.org/10.1111/1469-8986.3740551

Voss, J. L., Schendan, H. E., & Paller, K. A. (2010). Finding meaning in novel geometric shapes influences electrophysiological correlates of repetition and dissociates perceptual and conceptual priming. *NeuroImage*, *49*(3), 2879–2889. https://doi.org/10.1016/j.neuroimage.2009.09.012

West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research*, *13*(3), 363–375. https://doi.org/10.1016/S0926-6410(01)00129-X

Whitson, L. R., Karayanidis, F., Fulham, R., Provost, A., Michie, P. T., Heathcote, A., & Hsieh, S. (2014). Reactive control processes contributing to residual switch cost and mixing cost across the adult lifespan. *Frontiers in Psychology*, *5*. https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00383

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of Concurrent Tasks: A Psychophysiological Analysis of the Reciprocity of Information-Processing Resources. *Science*, *221*(4615), 1080–1082. https://doi.org/10.1126/science.6879207

Wodniecka, Z., Casado, A., Kałamała, P., Marecka, M., Timmer, K., & Wolna, A. (2020). Chapter Eight—The dynamics of language experience and how it affects language and cognition. In K. D. Federmeier & H.-W. Huang (Eds.), *Psychology of Learning and Motivation* (Vol. 72, pp. 235–281). Academic Press. https://doi.org/10.1016/bs.plm.2020.02.005

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. https://doi.org/10.1016/j.neuropsychologia.2020.107390

Wolff, N., Mückschel, M., Ziemssen, T., & Beste, C. (2018). The role of phasic norepinephrine modulations during task switching: Evidence for specific effects in parietal areas. *Brain Structure and Function*, *223*(2), 925–940. https://doi.org/10.1007/s00429-017-1531-y

Wolna, A., Łuniewska, M., Haman, E., & Wodniecka, Z. (2022). Polish norms for a set of colored drawings of 168 objects and 146 actions with predictors of naming performance. *Behavior Research Methods*. https://doi.org/10.3758/s13428-022-01923-3

Wong, A. S. W., Cooper, P. S., Conley, A. C., McKewen, M., Fulham, W. R., Michie, P. T., & Karayanidis, F. (2018). Event-Related Potential Responses to Task Switching Are Sensitive to Choice of Spatial Filter. *Frontiers in Neuroscience*, *12*. https://www.frontiersin.org/articles/10.3389/fnins.2018.00143

Wylie, G. R., Murray, M. M., Javitt, D. C., & Foxe, J. J. (2009). Distinct Neurophysiological Mechanisms Mediate Mixing Costs and Switch Costs. *Journal of Cognitive Neuroscience*, *21*(1), 105–118. https://doi.org/10.1162/jocn.2009.21009

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. https://doi.org/10.1016/j.bandl.2021.105043

Zhang, H., Kang, C., Wu, Y., Ma, F., & Guo, T. (2015). Improving proactive control with training on language switching in bilinguals: *NeuroReport*, *26*(6), 354–359. https://doi.org/10.1097/WNR.0000000000000353

Zhuo, B., Zhu, M., Cao, B., & Li, F. (2021). More change in task repetition, less cost in task switching: Behavioral and event-related potential evidence. *European Journal of Neuroscience*, *53*(8), 2553–2566. https://doi.org/10.1111/ejn.15113

8 Appendix

8.1 Appendix Investigation 1

8.1.1 General model including the factor word-age of acquisition

Table 4

Fixed effects for the LME model for naming latencies of the general model including the factors lexical word-frequency and word-age of acquisition.

Effect	Estimate	SE	t	by-Picture SD	by-Participant SD
Intercept	-1.10	0.04	-25.62***	0.15	0.22
Group	-0.07	0.05	-1.50		
Context	-0.04	0.04	-0.99		0.24
Word-lexical frequency	-0.01	0.01	-0.56	-	
Word-age of acquisition	-0.01	0.01	-0.41	-	
Age	-0.02	0.02	-1.05		
Age of acquisition	0.02	0.02	1.14		
log (Trial number)	0.01	0.01	1.13		0.03
Group:Context	0.08	0.06	1.45		
Group:Word-lexical frequency	-0.01	0.01	-1.29	-	
Group:Word-age of acquisition	0.00	0.01	0.04	-	
Control Group:Context:Word-lexical frequency	0.00	0.01	-0.13	-	-
Mig. Group:Context:Word-lexical frequency	0.02	0.01	2.63**	-	-
Control Group:Context:Word-age of acquisition	0.00	0.01	-0.01	-	-
Mig. Group:Context:Word-age of acquisition	0.00	0.01	0.18	-	-

Note. Because word-age of acquisition was not available for all items only 151 out of 216 items were used in this model.

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

8.1.2 Migrant group model including the factors length of L2 immersion (years living in the L2-environment) and length of L1 re-immersion (days spent in the L1-environment)

Table 5

Fixed effects for the LME model for naming latencies of the migrant group model with length of L1 reimmersion and L2 immersion.

Effect	Estimate	SE	t	by-Picture SD	by-Participant SD
Intercept	-1.14	0.04	-30.62***	0.16	0.15
Context	0.04	0.02	2.33*	-0.17	-0.18
Word-lexical frequency	0.00	0.01	0.05	-	
Length of L2 immersion	0.04	0.03	1.31		-
Length of L1 reimmersion	0.01	0.02	0.39		-
Age	0.00	0.03	0.01	0.78	-
Age of acquisition	0.01	0.02	0.59		-
log (Trial number)	0.00	0.01	0.41		-0.50
Context:Word-lexical frequency	0.01	0.01	1.67'	-	
Context:Length of L2 immersion	0.00	0.02	0.14		-
Context:Length of L1 reimmersion	-0.01	0.02	-0.42		-

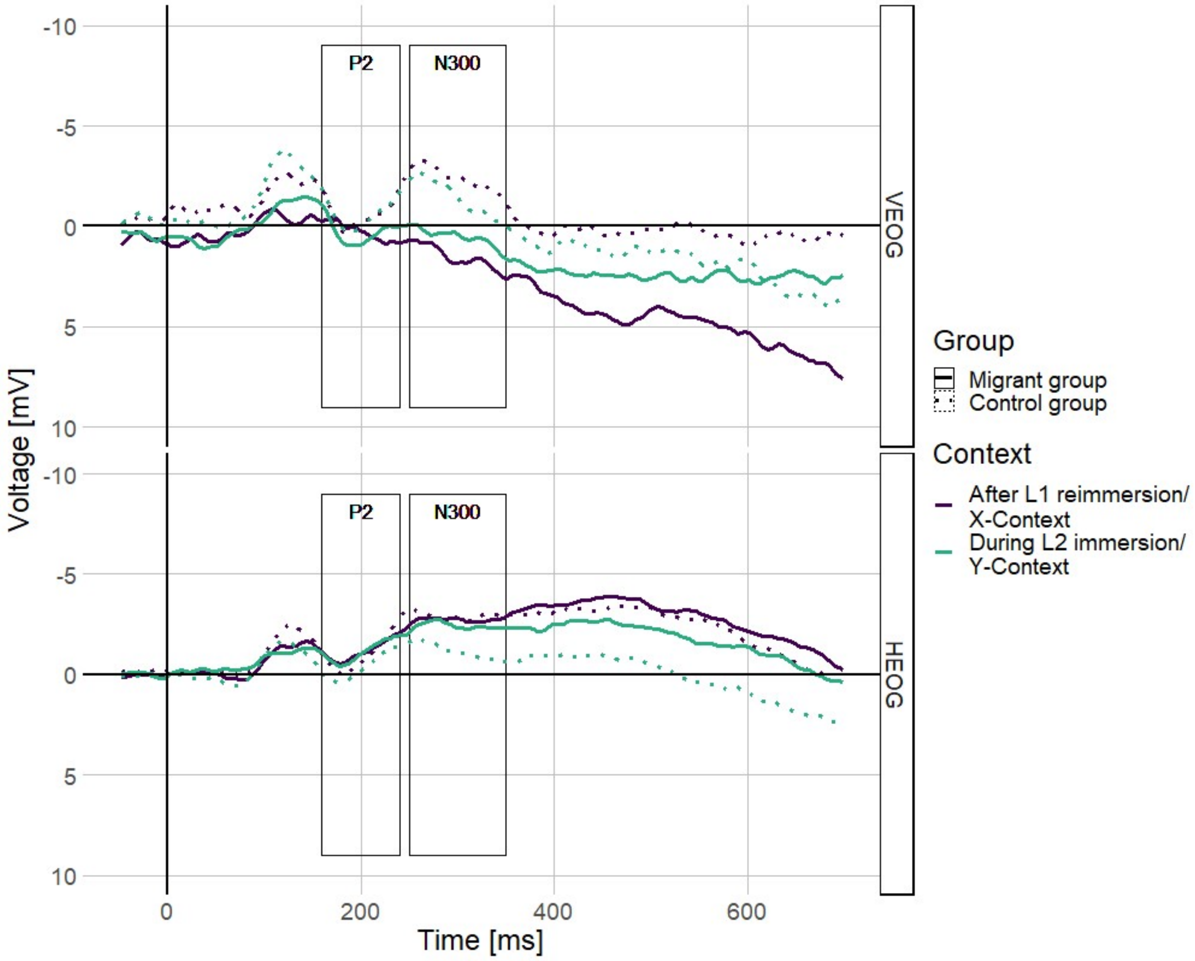
Note. Length of L2 immersion describes how many years a participant of the migrant group lived in the L2 environment. Length of L1 reimmersion describes how many days a participant of the migrant group stayed in the L1 environment, immediately prior to the experimental session.

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

8.2 Appendix Investigation 2

In order to discount the influence of ocular movements on the results, we investigated average eye blinks across groups and contexts. While the mean amplitudes of the ocular electrodes diverge in later time windows, there was no effect on the context or group effects in our selected time windows.

Figure 2.3



Note. Stimulus-locked grand average ERP waveforms for vertical (VEOG) and horizontal (HEOG) eye movements. Waveforms depict differences in mean voltages recorded from electrodes placed below and above a participant's left eye and at the outer canthus of each eye, respectively. separate electrode clusters for each component. Rectangles indicate time windows used to measure mean ERP amplitudes tracking the temporal dynamics of the target components: P2 (160–240 ms) and N300 (250–350 ms).

8.3 Appendix Investigation 3

8.3.1 Prolific questionnaire

Private users of Prolific can fill out several questions related to their background. Researchers can invite participants according to those answers and create their own research sample of the population. We used the following questions and answers to filter for our research:

1. What is your date of birth?
    a. Age 20 to including 35
2. Are you left or right-handed?
    a. Right-handed
    b. Ambidextrous
3. What is your nationality?
    a. Poland
4. What is your first language?
    a. Polish
5. Apart from your native language, do you speak any other languages fluently?
    a. native language + one other language
    b. native language + two other languages
    c. native language + three or more other languages
6. Which of the following languages are you fluent in?
    a. Polish
    b. English

We also excluded any participants that had already taken part in our previous pilot experiments.

8.3.2 Participant language proficiency of Pilot 1 and Pilot 2

Table 3.10 Self-assessed language proficiency of participants.

	Pilot 1 Participants		Pilot 2 Participants	
	Polish	English	Polish	English
Speaking	8.45 (1.77)	6.58 (1.67)	8.75 (1.49)	6.78 (1.76)
Writing	8.31 (1.32)	7.31 (1.55)	8.80 (1.53)	7.15 (1.71)
Listening	8.86 (1.59)	7.66 (1.46)	9.20 (1.38)	7.80 (1.71)
Reading	9.31 (0.98)	8.46 (1.24)	9.27 (1.29)	8.53 (1.35)

Note. Responses to the self-assessment proficiency questionnaire were on a 10-point scale, with 1 being the lowest and 10 the highest (native-like) proficiency. The presented values are mean response with standard deviations in brackets.