Evidence of non-selective lexical access to second and third language in unbalanced multilinguals: an N400 study on the processing of interlingual homographs

http://dx.doi.org./10.11606/1982-8837264935

Marina Fernandes Neves Lameira[1]

Felipe Rodrigues Bezerra[2]

Pâmela Freitas Pereira Toassi[3]

André Mascioli Cravo[4]

Maria Teresa Carthery-Goulart[5]

Abstract: Bilinguals and multilinguals commonly encounter words in their multiple languages which share some linguistic aspects. Among those are interlingual homographs, or words that have the exact same orthography in two different languages. The current study examined, through a semantic judgment task in English with ERP recording, how multilinguals (speakers of Brazilian Portuguese, English and German) accessed the meanings of interlingual homographs that belonged to their dominant and non-dominant foreign languages compared to a control group of Brazilian Portuguese-English bilinguals. The findings demonstrated that multilinguals were slower to respond to the English-German interlingual homographs as compared to control stimuli (no homographs). The results also demonstrated that, when the interlingual homographs were

[1] Universidade Federal do ABC, Programa de Pós-graduação em Neurociência e Cognição, Centro de Matemática, Computação e Cognição (CMCC). Alameda da Universidade, s/n, Anchieta, São Bernardo do Campo, SP, 09606-045, Brasil. Email:lameira.marina@gmail.com. ORCID: 0003-1937-3408.

[2] Universidade Federal do ABC, Programa de Pós-graduação em Neurociência e Cognição, Centro de Matemática, Computação e Cognição (CMCC). Alameda da Universidade, s/n, Anchieta, São Bernardo do Campo, SP, 09606-045, Brasil. Email:f.bezerra@ufabc.edu.br. ORCID: 0002-2574-6488.

[3] Universidade Federal do Ceará, Departamento de Estudos da Língua Inglesa, suas Literaturas e Tradução; Programa de Pós-Graduação em Estudos da Tradução (POET) e Programa de Pós-Graduação em Linguística (PPGLin). Av. da Universidade, 2853, Benfica, Fortaleza, CE, 60020-181, Brasil, Email: pamelatoassi@ufc.gmail.com. ORCID: 0003-3273-639X.

[4] Universidade Federal do ABC, Centro de Matemática, Computação e Cognição (CMCC); Programa de Pós-graduação em Neurociência e Cognição. Alameda da Universidade, s/n, Anchieta, São Bernardo do Campo, SP, 09606-045, Brasil. Email:andre.cravo@ufabc.edu.br. ORCID: 0002-8580-5697.

[5] Universidade Federal do ABC, Centro de Matemática, Computação e Cognição (CMCC); Programa de Pós-graduação em Neurociência e Cognição. Alameda da Universidade, s/n, Anchieta, São Bernardo do Campo, SP, 09606-045, Brasil. Email:teresa.carthery@ufabc.edu.br. ORCID: 0002-2751-4541.

Research supported by FAPESP, Grant Number 2017/00246-0 and CAPES (001).

semantically related to their targets in the non-target language, there were significantly more errors and a higher RT than in unrelated conditions. Additionally, only the bilinguals presented the typical N400 effect for unrelated conditions, suggesting that the co-activation of the non-target language due to interlingual homographs modulated this ERP in the multilingual group. Our results provide support for the Bilingual Interactive Activation plus model and suggest that literature findings on interference between first and second languages also hold for second and third languages.
Keywords: multilingualism; interlingual homographs; N400; Portuguese-English-German multilingualism

Zusammenfassung: Zweisprachige und mehrsprachige Menschen stoßen in ihren verschiedenen Sprachen häufig auf Wörter, die einige linguistische Aspekte gemeinsam haben. Dazu gehören interlingualen Homographen, d.h. Wörter, die sich in zwei unterschiedlichen Sprachen durch genau dieselbe Orthographie auszeichnen. In der vorliegenden Studie wurde anhand einer semantischen Beurteilungsaufgabe in Englisch mit ERP-Aufzeichnung untersucht, wie mehrsprachige Personen (Sprecher des brasilianischen Portugiesisch, des Englischen und des Deutschen) im Vergleich zu einer Kontrollgruppe von brasilianischen Portugiesisch-Englisch-Zweisprachigen auf die Bedeutung von interlingualen Homographen zugreifen, die zu ihrer dominanten und nicht-dominanten Fremdsprache gehören. Die Ergebnisse zeigten, dass mehrsprachige Personen im Vergleich zu den Kontrollstimuli (keine Homographen) langsamer auf die interlingualen Homographen Englisch-Deutsch reagierten. Die Ergebnisse zeigten auch, dass, wenn die interlingualen Homographen semantisch mit ihren Zielen in der Nicht-Zielsprache verwandt waren, signifikant mehr Fehler und eine höhere RT auftraten als in nicht verwandten Bedingungen. Darüber hinaus zeigten nur die Bilingualen den typischen N400-Effekt für nicht verwandte Bedingungen, was darauf hindeutet, dass die Koaktivierung der Nicht-Zielsprache durch interlinguale Homographen diese ERP in der mehrsprachigen Gruppe modulierte. Unsere Ergebnisse stützen das *Bilingual Interactive Activation plus* Modell und legen nahe, dass die Literaturergebnisse zur Interferenz zwischen Erst- und Zweitsprache auch für die Zweit- und Drittsprache gelten.
Stichwörter: Mehrsprachigkeit; interlinguale Homographe; N400; Portugiesisch-Englisch-Deutsch Mehrsprachigkeit

1 Introduction

1.1 The processing of interlingual homographs by multilinguals

There is a large debate on the neural organization of the mental lexicon (KEMMERER 2014) and more complexity is added to this discussion when considering the influence of bilingualism and multilingualism (KROLL; DUSSIAS 2004). Experiments investigating lexico-semantic access to word representations have shed light on the nature of the bilingual lexicon and gathered consistent evidence on the nonselective lexical access both in alphabetic (LEMHÖFER; DIJKSTRA 2004; KERKHOFS *et al.* 2006; STUDNITZ; GREEN 2002; MACIZO; BAJO; MARTÍN 2010; NOVITSKIY *et al.* 2018) and logographic language systems (HSIEH *et al.* 2017).

In the context of this discussion, the Bilingual Interactive Activation plus Model (BIA+) (DIJKSTRA; VAN HEUVEN 2002) explains how bilinguals retrieve lexical and semantic knowledge from written words in non-sentential contexts (i.e. lexical decision tasks). According to BIA+, written input generates parallel activation of orthographic, phonological and semantic representations of all words an individual has learned (belonging to all languages he might speak). These representations are stored in a common lexical system and are co-activated, which means that they compete for selection. This idea has been supported by a great amount of research focusing on native and foreign languages (see WU; THIERRY 2010 for a review), or on first, second and third languages (LEMHÖFER; DIJKSTRA; MICHEL 2004).

Cognates and Interlingual Homographs (IHs) are the most employed stimuli to address the question of co-activation in bilinguals and multilinguals since they allow us to make inferences regarding the bilingual lexicon in a monolingual task. When orthography and meaning are shared, cognates (e.g.: the words 'hotel' and 'taxi' in many languages) faster reaction times (RT) in the recognition of these words (as compared to non-cognates) have been taken as evidence for co-activation and this has been known as "the cognate facilitation effect" (see HEREDIA; ALTARRIBA; CIEŚLICKA 2015 for a review). On the other hand, words with exact same orthography but different meanings in either language are named IHs (KERKHOFS *et al.* 2006) (e.g.: Tag, meaning 'day' in German).

The general hypothesis is that when bilinguals or multilinguals are using the contextually relevant language, cognates and IH are processed differently (faster or slower) than words belonging to only one of the languages they master. That is, cognates normally generate a facilitation effect and IHs can generate facilitation or conflict depending on the task (LEMHÖFER; DIJKSTRA 2004; VAN HEUVEN *et al.* 2008).

There are a great number of studies with bilingual samples evidencing the IH effect in different contexts, for example in word level (HOSHINO; THIERRY 2012), sentence level (ELSTON-GÜTTLER; GUNTER; KOTZ 2005), using prime words (KERKHOFS *et al.* 2006) and semantic incongruence (MACIZO; BAJO; MARTÍN 2010).

Examining the semantic incongruence approach, it was noted that the reading of an IH can be biased towards its foreign language meaning, when prime and target are related in the target language (KERKHOFS *et al.* 2006). However, the presence of a prime can also bias the reading of the IH towards its native language meaning, even though the foreign language is the target, generating interference (HOSHINO; THIERRY 2012). It is not clear whether this interference (of the strongest language) occurs between foreign languages of multilinguals as fewer studies have investigated co-activation and the assumptions of BIA+ model for multilinguals.

We found few studies with multilingual samples using cognates in word level (LEMHÖFER; DIJKSTRA; MICHEL 2004), one study using cognates in sentence level (AUTOR XXXX) and two using both cognates and IHs in word level (ZHU; MOK 2018; BARCELOS & ARÊAS DA LUZ FONTES 2021).

Lemhöfer, Dijkstra and Michel (2004) studied double and triple cognates processing in multilinguals. Dutch-English-German trilinguals performed a lexical decision task in their third language (German). They found that double and triple cognates (i.e. cognates which overlapped in Dutch and German but not English, or which overlapped in these three languages) were recognized faster than control stimuli (pure German words). A greater facilitation effect for triple (in L1- L2-L3) than double cognates was found.

Zhu and Mok (2018) conducted two experiments using a lexical decision task with double cognates and one experiment with IHs between both foreign languages (L2 English and L3 German) of Cantonese-English-German trilingual individuals. The results showed both cognate facilitation (experiments 1 and 3), cognate conflict (experiment 2) and a small homograph facilitation (non-significant) in experiment 3 affecting the accuracy but not the reaction times. These distinct results patterns probably occur according to the target language of the task (L2 in experiment 1 and L3 in experiments 2 and 3) and also the lower frequency of IHs words and low L3 participant's proficiency. These factors could make the participant access the stronger meaning in L2, inducing the individual to answer in English instead German.

We also found two studies with Portuguese speaking multilinguals. The triple cognate facilitation effect was also evidenced in a sentence reading task with Brazilian-English-German multilinguals, in which trilinguals had their eye movements monitored while reading sentences in English containing cognates in their three languages (TOASSI et al. 2020). The cognate facilitation effect was found only for early comprehension measures: first pass and first fixation for the trilingual speakers. No effect for the triple cognates was found for the bilingual group, formed by speakers of Brazilian Portuguese and English and for the control group formed by native speakers of English.

Barcelos and Arêas da Luz Fontes presented their Portuguese-English-French trilinguals with a wide range of language proficiency a lexical decision task in their L3 French. Stimuli included double and triple cognates and IH. Participants had a higher accuracy rate with cognate stimuli compared to controls, but reaction times have shown no statistically significant difference to controls. IH yielded more errors, even though their effect was a marginal one, and RTs did not differ from controls.

While behavioral studies have succeeded in showing overt responses to the processing of IH (STUDNITZ; GREEN 2002; DIJKSTRA *et al.* 2000; LEMHÖFER; DIJKSTRA 2004; DE BRUIJN *et al.* 2001; GROOT; DELMAAR; LUPKER 2000; MACIZO; BAJO; MARTÍN 2010), neuroscientific methods have evolved to analyze the neural mechanisms underlying cognitive processes, known as endogenous components. One of these methods is electrophysiology, particularly the Event-Related Potentials (ERP) technique (KUTAS; PETTEN; KLUENDER 2006) that allows the measurement of the N400 component, whose amplitude is modulated by lexical-semantic processing (KERKHOFS *et al.* 2006; HOSHINO; THIERRY 2012; KOUSAIE; PHILLIPS 2011) and interlingual incongruence (THIERRY; WU 2004; THIERRY; WU 2007).

In language experiments, the N400 component indexes semantic integration processes, presenting a negative amplitude 400 ms after the presentation of a pair of items. When subjects are confronted with unexpected semantic violations, the amplitude of the N400 component is more negative than the amplitude elicited by items with no semantic violation (KUTAS; FEDERMEIER 2011). In neutral contexts, words which occur more frequently in the natural speech of a certain language and words which more familiar to individual speakers elicit smaller and less negative N400 amplitudes. Kerkhofs *et al.* (2006) investigated the effect of primes as a bias towards the low-frequency, but task-relevant reading of an IH by analyzing N400 amplitudes. Frequent words in L1, which should elicit a less negative amplitude than low-frequency words in L2, presented more negative amplitudes when their meaning was irrelevant to the task. The finding corroborated the BIA+ model, which postulates that the selection of lexical representations is sensitive to the task demands.

Few studies approached whether a priming effect towards the non-target language could be reflected on N400 amplitudes. Hoshino and Thierry (2012) addressed this issue with Spanish-English bilinguals and English monolingual controls in a semantic judgment task. They compared the N400 mean amplitude for word pairs which were unrelated in English (L2) but related in Spanish (L1) and found a modulation of the N400 resulting from the presentation of a prime in Spanish (non-target language).

Similarly to Dutch, German shares a good number of IHs with English, providing enough resources to assess the co-activation effect in this population of speakers. However, few studies explored this language combination. For example, Von Studnitz and Green (2002) conducted a lexical decision task with English-German homographs, and Elston-Güttler (2005) did ERP work with English-German homographs in sentence level, both showing an interference effect of IHs on semantic processing.

We have no knowledge of studies examining the semantic lexical access in multilinguals using both EEG and behavioral measures. Because of this gap, in the present study, we combined these approaches to examine if the IH effect can modulate the semantic processing between two foreign languages. In other words, we aimed at investigating whether the knowledge of a second foreign language should generate a different wave pattern for multilinguals in comparison to bilinguals when faced with a word with different meanings across these two languages.

1.2 The current study

Our main objective was to investigate whether the assumptions of the BIA+ Model could be further extended to multilinguals. The study compared Portuguese-English bilinguals to Portuguese-English-German multilinguals living in a Portuguese-speaking context. The bilingual group (our control group) was fluent in English and the multilingual group (our experimental group) was fluent in English and German. For the multilingual group, co-activation of the two foreign languages (L2-L3) was analyzed by assessing the possible interference of the dominant foreign language (L3- German) in their non-dominant foreign language (L2-English). As previous studies have found an interference of the native language (L1) in the non-native (L2) language, we were interested in investigating whether these findings should also apply to the foreign languages of unbalanced multilinguals (hence L2 and L3). Our hypothesis was that multilinguals would suffer an interference from their dominant foreign language (German) into the non-dominant foreign language (English). We avoided the overlap with Portuguese to the extent that we did not include words with a Latin origin.

In the semantic judgment task in English, pairs of English words which were related or unrelated in meaning were presented. We call these “related” (required a "yes" response) and “unrelated” (required a "no" response) conditions. Half of the stimuli consisted of word pairs without IH which could be either related (e.g.: tree – leaf) or unrelated in English (e.g.: tree – pan). These were the control conditions, as they were not expected to induce any conflict for German speaking participants. Our experimental conditions, on the other hand, had English-German IH, i.e. English words which also exist in the German language with exact same orthography but have different meanings (e.g.: boot – boat in German). Relatedness of these pairs considering the German (non-target) reading of the IH was opposite of their English (target) reading, e.g.: boot – foot: unrelated in German but related in English, and boot – lake: related in English but unrelated in German. Given that our stimuli in the experimental condition consisted of word pairs which were unrelated in English but related in German, or related in English and unrelated in German, we hypothesized that these conditions would result in lower accuracy rates and longer RTs compared to control conditions (without IHs) in the multilingual group. We expected that our control group (bilinguals) would on the other hand not suffer this influence of IHs.

We measured these effects in terms of the overt behavioral response (reaction time – RT– and response accuracy) and covert processes indexed by the N400 component. The experimental conditions aimed at inducing a conflict on the selection of the target language (L2- English). Furthermore, we intended to see whether the N400 component is modulated by whether the individuals are bilinguals or multilinguals. In other words, we expected bilinguals and multilinguals to exhibit different amplitude patterns in the N400 time window due to different activations of semantic relationships. As previous studies have shown, unrelated conditions elicit higher amplitudes of the N400 component. Both groups were therefore expected to exhibit higher amplitudes of the N400 component in the control unrelated conditions. However, similarly to the study of Hoshino and Thierry (2012) with bilinguals only, experimental conditions with IHs should induce a conflict in the multilingual group, reflected by amplitude patterns different from the ones exhibited by the bilinguals.

2 Method

2.1 Participants

The participants of this study were 24 native speakers of Brazilian Portuguese, divided in the following groups: (1) the multilingual group was formed by 13 participants (6 males, mean age 29.5 years, SD: 6.6) who had German as their dominant foreign language (L2), being also fluent in English (L3); (2) the bilingual group was formed by 11 participants (4 males, mean age 29.2 years, SD: 6.6) who had English as their dominant foreign language and no knowledge of German. Proficiency in the participants' foreign language was assessed by objective language tests (see Appendix). The Research Ethics Committee of the Federal University of ABC approved the experimental protocol (Nr. 2.182.943).

Regarding foreign language acquisition, 91.66% of the bilingual participants do not consider themselves to speak another language at an advanced level, except for one individual who reported speaking Spanish. Concerning immersion experience, only 25% of the bilingual participants reported having lived in an English-speaking country for a period no longer than 1 year. As for the multilingual participants, 81.8%reported having acquired English as their first foreign language, whereas 9.1% acquired German first and 9.1% German and English simultaneously. Most of them (81.8%) consider German to be their dominant foreign language (18.2% considered they were equally fluent in German and English). Concerning immersion experience, 81.8% reported having lived in a German speaking country, i.e., they have already been in an exogenous linguistic environment (BAKER, 2011) for a period not longer than 1 year. We considered subjects' language dominance according to their self-assessed current language use. Self-assessing questionnaires included their language acquisition through lifetime alongside with other measures, such as musical training and videogame practice. We used these data as supplementary information to our objective language measures and neuropsychological measures (for additional information on the subjects, see Appendix: characterization of the subjects). The objective language measures aimed at matching both groups according to proficiency in the task language, English, and the neuropsychological measures aimed at matching both groups according to fluid intelligence measures.

2.2 Description of the task and EEG recording

2.2.1 Semantic judgment task

The semantic judgment task was composed of nouns, presented in upper case letters to neutralize the distinction of nouns in German, which are written with a starting uppercase letter. Five from the selected words (2.78% of total) could be either a noun or a verb in English. The homograph nouns were selected according to word frequency based on SUBTLEX- US and SUBTLEX-DE (BRYSBAERT; NEW 2009; BRYSBAERT ET AL 2011) databases for the English and German reading of the homographs, respectively. For all primes and target pairs, we composed the initial selection of candidates for the stimuli list using the following synonym and thesaurus tools: Macmillan Online (English) (MACMILLAN, 2017), DWDS (WISSENSCHAFTEN 2017) and Duden Online (DUDEN 2017) (German).

2.2.1.1 Task stimuli

The four conditions totalized 114-word pairs (171 English words, being 29 interlingual homograph primes and 29 related targets, the same 29 interlingual homograph primes and different 29 unrelated targets, 28 non-homograph control primes and 28 related targets, the same 28 non-homograph control primes and different 28 unrelated targets. One pair of non-homograph controls was eliminated due to low familiarity found in the pilot phase. These word pairs were divided into two experimental blocks of 57 trials each. In addition, a practice block of 29-word pairs which preceded the experimental blocks was not analyzed. Primes were either 29 identical German/English IH or 28 English control words paired with target non-homograph words. The primes were presented randomly in the related and unrelated conditions. Target words were not repeated in the experiment. IH were orthographically identical in German and English (100% spelling overlap) and all words were nouns – although 2,78% could also be read as verbs. IH with Latin etymology were excluded, in order to avoid interference from the native language, Portuguese, which is a romance language. Examples of the four experimental conditions and details of the task concerning the words pairs are shown in Figure 1. The psycholinguistic features of the stimuli and the preliminary study to create the task are described in the appendix.

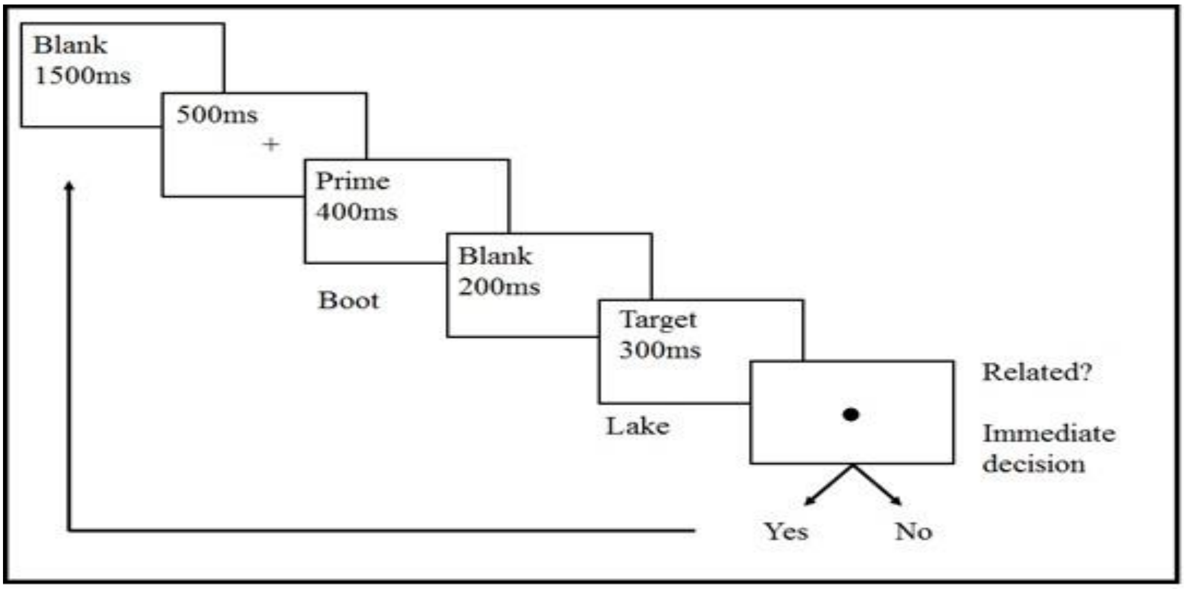
2.2.1.2 Procedure

For the semantic judgment task, participants were seated in an electrically shielded chamber in front of a 17 inches screen and adjustments were made for the cap placement, as well as signal recording. Stimuli were presented with E-Prime Software Version 2.0. During these procedures, participants were asked to keep English language mode in interactions with the experimenter.

The practice block (29 trials) was followed by two experimental blocks (57 trials each) with an interval between them. Each trial began with a fixation cross that remained on screen for 500ms, followed by the prime for 400ms, a blank screen interval for 200 ms, the presentation of the target for 300ms, and a following screen with a dot, which remained on screen until the participant pressed either the key 1 or 9 in the answer box to judge whether the word pair was related or unrelated. Answer keys were counterbalanced, so that half of the participants in the same group pressed the right button for "yes" and the left button for "no", and the other half, the opposite way around. After the participant's response, a blank screen preceded the following trial for 1500ms (see figure 1). Words were presented in black on a white background.

After the experiment, participants completed a post-test in which they had to translate to Portuguese all words from the stimuli list (English to Portuguese translation) without consulting dictionaries or any supplementary materials. A minimum of 70% accuracy was required for the participant to be included in the analyses. Mean accuracy and medianRT of participants in both groups was calculated on E-Prime Software Version 2.0–Professional. We chose the median as it is a more robust measure compared to the mean, given that the mean can be affected by extreme values, especially with a modest N, which was the case. A repeated measures ANOVA was performed separately for accuracy and RT rates, with "relatedness" and "homograph" as within-subject factors and group as between-subjects' factors. Participants performed the task while electrophysiological data were recorded.

Figure 1: Semantic Judgment task description and examples of conditions



Conditions	Example of a pair for semantic judgement	Related (English)	Related (German)	Homograph (prime)
Control condition: Related without IH	a) TREE x LEAF	+	n/a	-
Control condition: Unrelated without IH	b) TREE x PAN	-	n/a	-
Experimental condition: Related with IH	c) BOOT x FOOT	+	-	+
Unrelated with IH	d) BOOT x LAKE	-	+	+

Source: Own authorship

2.2.2 EEG recording

EEG data were collected with use of an elastic cap with 64 active electrodes (actiCAP – Brain Products GmbH, Gilching, Germany), positioned in international system 10-10. Two additional bipolar electrodes were used to register oculogram. Vertical electrodes (VEOG) were positioned above and below the right eye and horizontal electrodes were positioned to the left and the right of the eye line. EEG was continuously recorded at a 1000 Hz rate with a QuickAmp amplifier. Ground (Afz) and reference (FCz) electrodes were re-referenced to the earlobes. Parallel port fed the markers of the stimuli from the stimulation computer into the amplifier and recorded within the EEG data set.

2.2.2.2 EEG preprocessing and analysis procedures

For the EEG analysis, data from 19 participants were included (10 from the multilingual group and 9 from the bilingual group). Data from 3 participants were removed due to technical problems during EEG collection and data from 2 participants were removed due to excessive noise in the EEG signal. EEG data were analyzed offline with Brain Vision Analyzer 2.1 software (Brain Products, Germany). EEG data was downsampled and to 250Hz and filtered (low-pass at 30Hz, 24dB, high-pass at 0.05Hz, 24dB, and a notch filter at 60Hz). Presence of flat or excessively noisy channels was evaluated, with use of spherical spline interpolation method, if necessary.

In sequence, data were separated in epochs of 1350 ms, with 750ms before target onset and 600 ms after target onset. The period before target onset includes the last 150ms of the fixation cross (150ms), the presentation of the prime (400ms) and the black before target presentation (200ms). For eye movement artifact rejection, an independent component analysis (ICA) was performed and eye-related components were identified semi-automatically and by visual inspection of topographies and time series from each component. Eye related components were then rejected for all segments. Baseline correction was performed using the whole pre-target period. Trials that exceed 150 µV were rejected (mean: 5%, max of 29% of trials removed due to artifacts).

Sensors and moments of interest were chosen based on the grand-averaged N400, which was calculated by subtracting the EEG activity of all unrelated targets from all related targets, across all conditions and groups. In order to compare N400 mean amplitudes in the conditions with and without homographs for each group, we subtracted the unrelated from the related conditions with and without homographs and a repeated measures ANOVA was performed with homograph presence as within subjects factor, and group as between subjects factor.

3 Results

Relevant data will be uploaded on a website page of Open Science Framework. For further information, please contact authors.

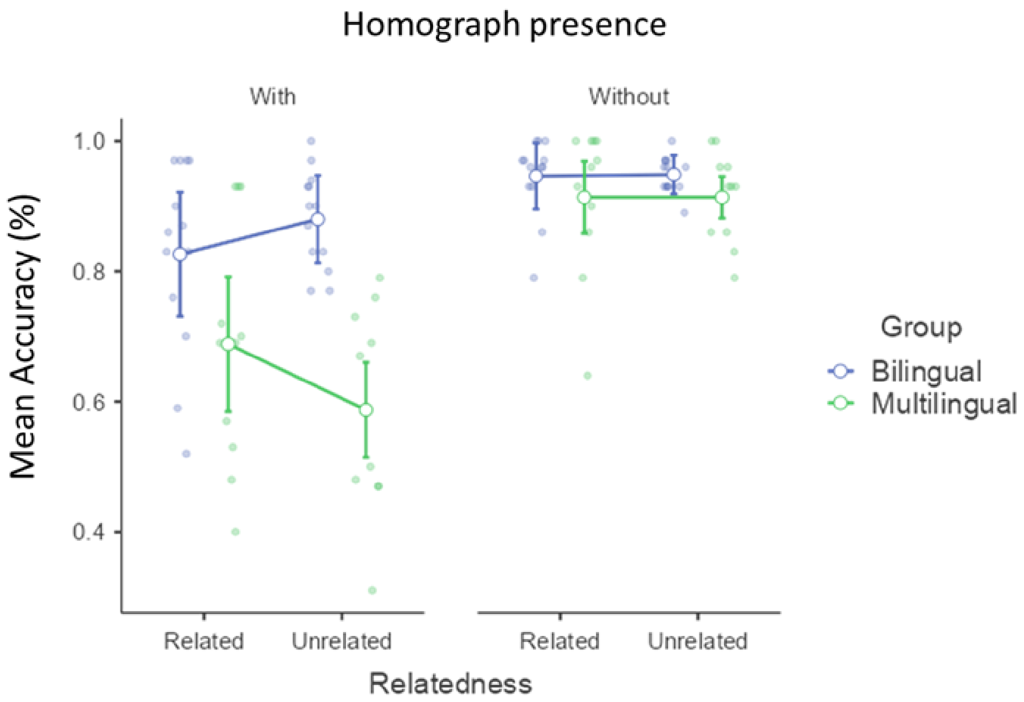
3.1 Behavioral results

3.1.1 Accuracy

Figure 2 shows the accuracy rates for both groups. Both groups had a mean accuracy rate of over 90% both in the related and unrelated conditions without homograph (bilingual - related: .94, unrelated .64; multilingual - related: .91, unrelated .91). In the conditions with homographs, the bilingual group had an average accuracy rate of over 80% (related: .82, unrelated: .88), while the multilingual group scored lower (related: .68, unrelated: .58).

As the participants showed high accuracy rates, we used a non-parametric test, Mann-Whitney test (which compares differences between two independent groups). Specifically, we compared pairwise performances between groups for each of the four conditions (related with and without homograph, and unrelated with and without homograph) and corrected the observed p-values using Bonferroni correction to adjust critical p-value according to the four comparisons, which then was set to .012.

Figure 2: Semantic relatedness judgement task: Mean Accuracy rates per condition and group



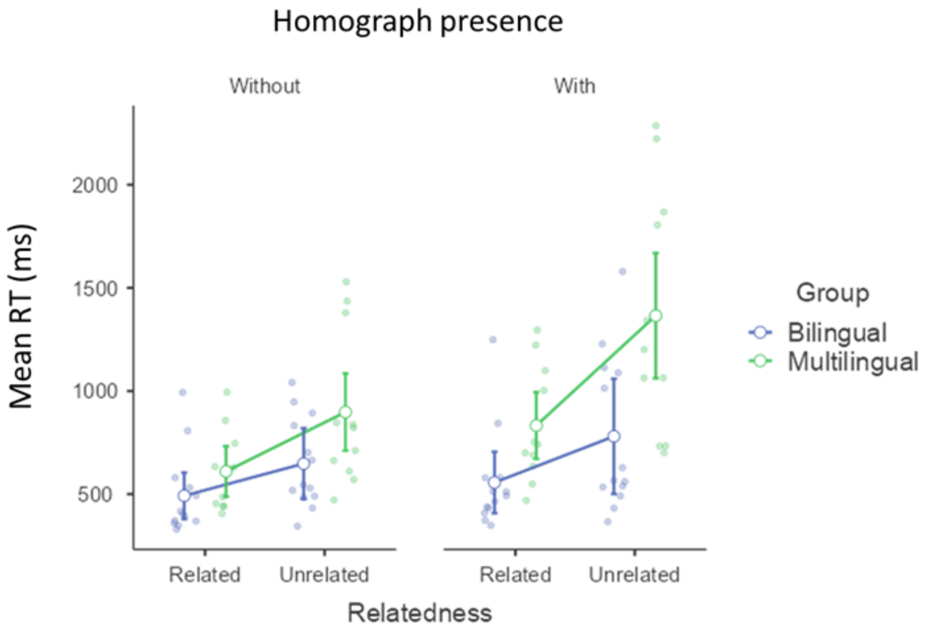
Source: Own authorship

We found a significant effect between groups only in the unrelated with homograph (unrelated in English/related in German) (critical) condition ($p < .0001$), where the multilingual group had lower accuracy (mean = .58 SD = .15) than the bilingual group (mean = .88 SD = .07) with large effect size (rpb = .97). No other differences were significant (all corrected p-values > .23). These results indicate that the presence of a homograph affected more the group who speaks two foreign languages (the multilingual group), generating a conflict between the two word representations, i.e., the English reading and the German reading of the IH.

3.1.2 Reaction Time (RT)

We calculated RT for the correct responses only. In order to exclude extreme scores, RT median for each subject was used. A repeated-measures ANOVA showed significant main effects for a) group ($F= 6.90$, $p= .015$, $\eta^2p= .239$): bilinguals answered in general faster than multilinguals; b) relatedness ($F= 35.02$, $p < .001$, $\eta^2p= .614$) showing higher RTs in unrelated conditions for both groups and c) homograph ($F= 40.66$, $p < .001$, $\eta^2p= .649$) evidencing higher RTs in homograph conditions for both groups.

Figure 3: Semantic relatedness judgment task: Mean reaction times per condition and group



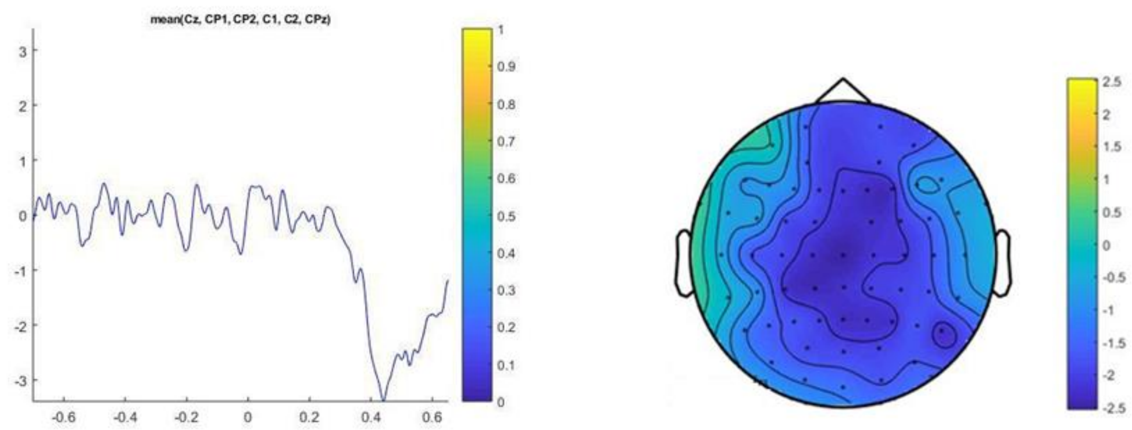
Source: Own authorship

There was also a significant interaction of relatedness*group (F= 4.73, p= .041, $\eta^2 p$= .177). To follow-up this interaction, we compared the cost in RT of the word being unrelated for each group. We found that the cost was stronger for multilingual speakers than for bilingual speakers (t(22)=2.18, p= 0.041). A similar pattern was found for the interaction of homograph*group (F= 12.60, p= .002, $\eta^2 p$= .364). Similar to the previous analysis, we compared the cost of having a homograph word across groups. We found a higher cost in RTs for the multilingual group (Welch's t(16)= 3.41, p= 0.004). Lastly, there was an interaction of relatedness*homograph (F= 10.04, p= .004, $\eta^2 p$= .313), suggesting that the presence of the homograph had a stronger cost in unrelated than on related words. The three-way interaction was not significant (F= 3.23, p= 0.086, $\eta^2 p$= 0.128).

3.2 Electrophysiological results

Figure 4 shows the N400 in all conditions and groups and its correspondent topographic map. On the left, N400 component conditions in electrodes Cz, CP1, CP2, CPz in all and groups. On the right, topographic map: the colors indicate mean amplitude in each region of the scalp.

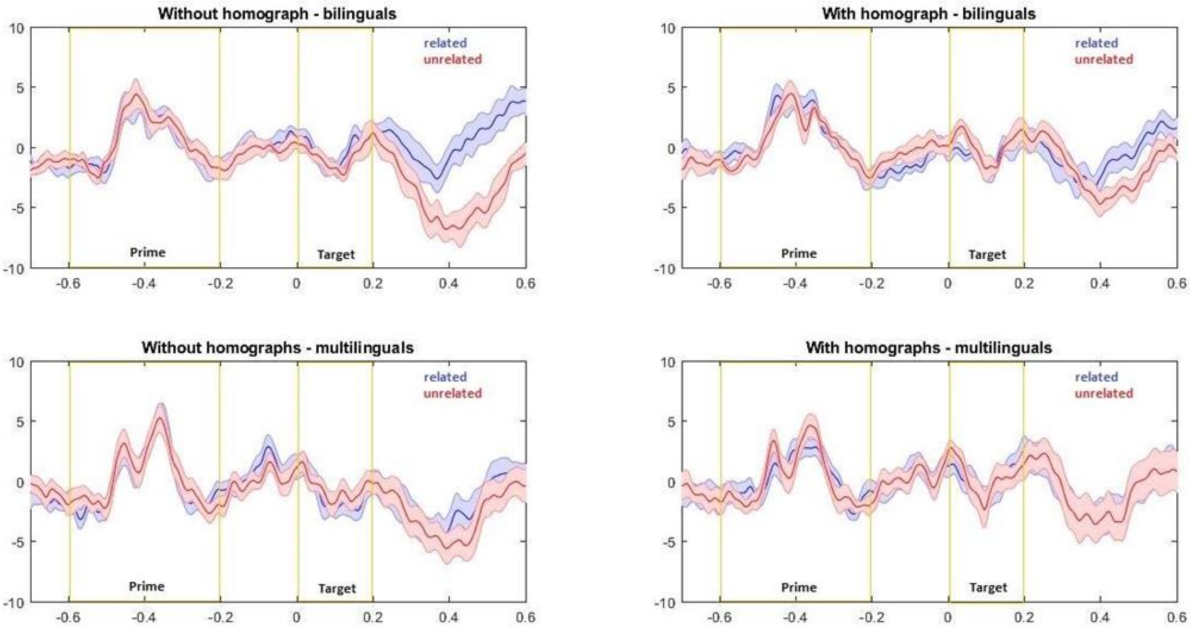
Figure 4: Electrophysiological results: N400 component (all conditions and groups) and topographic map.



Source: Own authorship

A visual inspection of the mean amplitude curves in the bilingual group (figure 5, top) indicates, in the 300-500ms time window, the common N400 finding with a more negative amplitude for unrelated words, with a smaller difference for words with homographs. On the other hand, the visual inspection of the multilingual (figure 5, bottom) group indicates a similar N400 across all conditions, irrespective of the words being related or unrelated or having or not homographs.

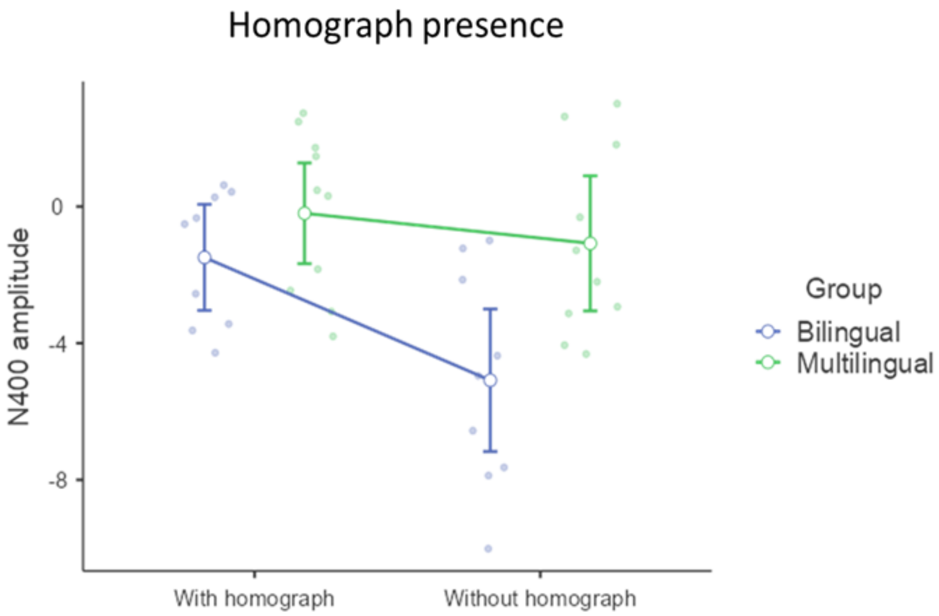
Figure 5: Electrophysiological results: Mean amplitude curves for the bilingual group and the multilingual group.



Source: Own authorship

In order to compare N400 mean amplitudes in the conditions with and without homographs for each group, we subtracted the unrelated from the related conditions with and without homographs, and a repeated-measures ANOVA was performed with homograph presence as within-subjects factor, and the group as between-subjects factor. We found significant effects for both main effects: homograph ($F= 8.81$, $p= .009$, $\eta^2p= .341$), with words without homographs eliciting a more negative n400; and group ($F= 8.05$, $p= .011$, $\eta^2p= .321$) with a more negative n400 for the bilingual group. The interaction homograph x group was not significant ($p= .09$, $\eta^2p= .160$). See figure 6 for mean subtracted amplitudes.

Figure 6: Electrophysiological results: Mean subtracted amplitudes for the bilingual and the multilingual groups



Source: Own authorship

4 Discussion

In this study, we tested the language non-selective view of lexical access proposed by the BIA+model applied to foreign languages only. Previous studies have shown that when bilinguals perform a task in a second language, there is co-activation of lexical items from the first language, which may compete for selection (STUDNITZ; GREEN 2002; LEMHÖFER; DIJKSTRA 2004; KERKHOFS *et al.* 2006; THIERRY; WU 2007; HEUVEN *et al.* 2008; MACIZO; BAJO; MARTÍN 2010; HOSHINO; THIERRY 2012).

To our knowledge, very few studies were performed with the aim of testing the BIA+ assumptions on foreign languages only. Based on findings that bilinguals suffer an interference from the native language into the foreign language and also on the other way around (HSIEH *et al.* 2017), we hypothesized that multilinguals would analogously suffer an interference from their dominant foreign language into their non-dominant foreign language. A recent study, with Cantonese- English- German trilinguals (ZHU; MOK 2018) presented evidence of this influence between foreign languages, in which there was an effect of the less dominant language – German – into the more dominant one - English. In the current study, in order to investigate whether the BIA+ model could be extended to foreign languages we studied a sample of unbalanced multilinguals, who considered

German as their dominant foreign language, and a sample of bilinguals, who considered English as their dominant foreign language.

The discussion is organized in two parts: firstly, we discuss our results (behavioral and electrophysiological) in light of the literature; secondly, we discuss our main findings considering the language non-selective view.

4.1 Behavioral results

Our study showed: a) a higher accuracy rate for conditions without homographs in both groups; b) a higher general accuracy rate in the bilingual group; c) a much lower accuracy rate in the multilingual group (58.73% of multilinguals against 87,58% of bilinguals), in the unrelated in English and related in German condition, i.e. the critical condition - the only condition which was unrelated in the target language (required a 'no' response, but was related in the non-target language (required the inhibition of a 'yes' response). Reaction times on correct responses indicated: a) multilinguals had higher RT than bilinguals in all conditions b) unrelated conditions and conditions with homographs had higher RTs for both groups and c) in the unrelated condition in English and related in German (critical), RTs were significantly higher for the multilingual group.

The slower responses for multilinguals compared to bilinguals may indicate that the more languages one speaks (or understands), the longer language selection takes. They are in line with the BIA+ Model, which proposes an integrated lexicon for bilinguals. Studies with bilinguals and control monolingual groups have shown that reaction times for bilinguals were longer than for monolinguals (GROOT; DELMAAR; LUPKER 2000; STUDNITZ; GREEN 2002; THIERRY; WU 2004; VAN HEUVEN *et al.* 2008). We hypothesize that the larger this lexicon becomes, more complexity arises during language processing. This finding was similar to the study by Durlik, Szewczy, Muszyński, and Wodniecka (2016), in which the interference of Polish-English interlingual homographs was also evident in a longer RT in a semantic relatedness judgment task.

In line with previous studies, unrelated conditions led to longer RT rates than related ones (THIERRY; WU 2004; THIERRY; WU 2007; KERKHOFS *et al.* 2006; KOUSAIE; PHILLIPS 2011). In the study of Thierry and Wu (2007), both groups had higher (and similar) accuracy rates in the related conditions. This result was also found in our study.

Conditions with IHs had longer RTs for both groups and more errors in the multilingual group than control conditions without IHs. The findings in the IH condition for the bilingual sample (despite they had no knowledge of German) were also seen in previous studies and is generally attributable to a difficulty in matching perfectly the conditions. IHs also impacted the performance in a monolingual condition in the study of Studnitz and Green (2002). The authors attributed this effect in part to differences in the subjective frequency of the stimuli (versus frequency obtained in corpora). They mentioned the complexity of creating experimental designs involving IHs and composing stimuli lists free of nuisance factors. Similarly, in our study, although stimuli were matched in frequency, familiarity ratings were significantly lower for conditions with IH, so we cannot completely rule out the impact of familiarity of items in the comparison of conditions with and without IH. Acknowledging this limitation, we observed significantly higher RTs and significantly lower accuracy in the multilingual group for the unrelated condition in English and the related condition in German. Our results corroborate previous findings on lexical decision studies (GROOT; DELMAAR; LUPKER 2000; HSIEH *et al.* 2017) and semantic relatedness studies (MACIZO; BAJO; MARTÍN 2010; THIERRY; WU 2004).

In the study of Kerkhofs et al. (2006), relatedness facilitated the selection of the correct reading of the IH: IH preceded by related primes were responded to faster than IH preceded by unrelated primes. In our study, IH were used in the prime position (similarly to the study of Macizo, Bajo and Martín, 2010). This design was preferred to allow for the two possible semantic representations of the IH to be accessed and to check then whether relatedness with the target, in the target language, could generate a disambiguation in the reading of the IH. Our results showed that relatedness in the target language facilitates the selection of the correct reading of the IH, given that conditions with IH related in English (and unrelated in German) were responded to faster than conditions related in German (and unrelated in English). Relatedness led to facilitation in the control conditions without IHs and disambiguation in the conditions with IHs, whereas unrelatedness led to conflict in the multilingual group. The critical condition (unrelated in English and related in German) had significantly higher RT and error rates, suggesting that the strong association between the non-target reading of the IH and its prime could either not be inhibited at all, leading to incorrect responses, or demanded a greater effort to be inhibited, leading to higher RTs. In line with studies of IH in sentence context (ELSTON-GÜTTLER; GUNTER; KOTZ 2005), these results indicate a context-dependent lexical selection phase: there was only disambiguation on the IH selection when its English reading was related to the target.

The study of Van Heuven and colleagues (2008) found evidence of conflict in the bilingual selection of the IH only in RT and not in accuracy rates. Here we point out possible explanations for this discrepancy with our study: first, their bilinguals were foreign students living and studying in an English-speaking country, with immersion in the L2 language and culture. Our sample of bilinguals and multilinguals, in turn, were living and studying in their homeland and immersed in a different language (Portuguese). Even though some of them may practice their L2 and L3 at work, they were not immersed in a foreign language context. Due to this fact, participants from the study of Van Heaven and colleagues (2008) may have been less susceptible to interference from their strongest language, since they were using their weakest language daily. Second, as first and second languages may have different representations in the brain, depending on the individual's relative proficiency between these languages, and foreign languages may have to a greater or lesser degree overlapped representations, it may be easier to inhibit the native language rather than a stronger foreign language while dealing with the weaker one.

Another unexpected result is that error rates in the critical conditions in our study were greater than in previous studies. Hsieh and colleagues (2017), for example, tested whether Chinese-Japanese bilinguals would suffer interference from their second language, Japanese, in their first language, Chinese, during a lexical decision task. Interestingly enough, they found this interference in the L2-L1 direction, demonstrated in RTs only, while the accuracy rate remained at about 90% in all conditions. We argue here that this interference direction (the weakest language into the strongest language), which is the opposite of ours, may account for the high accuracy rate in that study. Besides that, as pointed out above, we cannot rule out the possibility that one's control between two foreign languages is weaker than between a native and a foreign language, which could explain our accuracy rates.

4.3 Electrophysiological results

As figure 4 indicates, by collapsing all conditions and groups, we found a typical N400 component in electrodes Cz, CP1, CP2, CPz. Homograph presence has affected mean amplitudes of both groups, with unrelated conditions without homographs more negative in amplitude in the bilingual group. In the multilingual group, there was no clear N400 effect, once all conditions have elicited a similar amplitude pattern. Studies with semantic judgment comparing monolinguals with bilinguals evidenced some trends which were also found in our work. In the studies by Thierry and Wu (2004) and Hoshino and Thierry (2012), their monolingual (control) group was shown paired-up related and unrelated conditions in the N400 time-window, with unrelated conditions presenting a more negative amplitude. In our study, this same pattern occurred in our bilingual (control) group. Also, in the study of Thierry and Wu (2007) and Kousaieand Phillips (2011) and in our study, the mean difference in amplitude between related and unrelated conditions was greater (and more evident) in the control group. These results are taken as evidence that the control group suffered no language conflict, as electrophysiological results demonstrate a clear target-reading of the IH, in our case and Kousaieand Phillips' (2011), and no translation to the non-target language, in the case of Thierry and Wu (2007).

In the study of Thierry and Wu (2007), monolingual and bilingual participants had to judge pairs which could be either related or unrelated in English (L2), whose translation equivalent to Chinese (L1) could have a repeated character (their incongruent condition, e.g.: *Huo Che* and *Huo Tui*, meaning train and ham) or not(their congruent condition). They found that monolingual speakers of Chinese had more difficulties judging unrelated Chinese pairs which had a repeated character than pairs which did not, and this result also occurred when bilingual speakers of Chinese and English had to judge English pairs with these characteristics, while monolingual speakers of English had no hard time judging the incongruent conditions at all. In the bilingual group, N400 lasted more and had a greater magnitude in the congruent conditions for bilinguals only, in contrast to incongruent conditions. In the study of Hoshino and Thierry (2012), however, the experimental group (Spanish-English bilinguals) showed similar processing patterns to the control group: related and unrelated conditions were clearly paired-up, but for monolinguals the related conditions were the ones whose English reading of the IH was related to its English prime and the other way around, while for bilinguals the related conditions were the ones whose Spanish reading of the IH was related to its English prime. Our multilingual (experimental) group, in turn, showed no clear and no statistical difference between related and unrelated conditions, or in control (without IH) or experimental (with IH) conditions. Even though our task was not the same as the task of Thierry and Wu (2007), in their bilingual group the ERP measure of congruent and incongruent conditions also seemed to be indicative of a less clear processing pattern, when compared to monolinguals. Our results are more in line with the findings of Thierry and Wu (2007) than Hoshino and Thierry's (2012).

4.4 General discussion: Evidence of non-selective access for foreign languages

Using a lexical decision task, Von Studnitz and Green (2002) found an external locus of control during the processing of IHs, i.e., bilinguals are able to reduce interference by controlling how they respond to them, but not by reducing the activation levels of the non-target language in their lexico-semantic system. They found a "carry-over effect", i.e., a delayed "yes" response in the lexical decision task, in trials preceded by an IH whose non-target reading generated residual activation in "no" units, did not dissipate when participants were informed on the type of conflict they would face during the experiment. The authors concluded (p. 17): "Conceivably, the carry-over effect simply reflects processing difficulty and carries no implications for the nature of control". We argue here that our behavioral and electrophysiological results provide support to this claim: ERPs showed that, for multilinguals, there was an initial phase in the processing of word recognition in which the co-activation was very strong and was spread even to control conditions, generating a great cognitive effort to interpret the words they were facing. RT reflected the size of the inhibition, whereas accuracy rates demonstrated a later phase of conflict resolution. These results are in line with the BIA+ Model (DIJKSTRA; VAN HEUVEN 2002).

Yiu and colleagues (2015) aimed at finding the temporal window for the language membership and semantic/syntactic information. Considering semantic priming effects from non-target language representations shown in previous research, they hypothesized that semantic and syntactic information should be accessed before language membership information. To investigate this issue, early balanced Spanish-English bilinguals performed a go/no go language task, in which they had to answer a semantic question (is this word an animal or an object) only on English words and withhold a response to the same question on Spanish words. Similarly, there was a go/no-go semantics task, in which participants had to respond to animals and objects in English and withhold responses to animals and objects in Spanish. Both conditions were counterbalanced between participants, with a final 8-instruction design. By evaluating the onset peak of N200, which reflects response inhibition, earlier language membership access was encountered, thus not confirming their hypothesis. On the other hand, in a Spanish Lexical Decision Task and Spanish-category decision task (answer only Spanish words referring to person), Ng and Wicha (2013) found proficiency-modulated results in comparison to the results of Yiu and colleagues (2015). Their participants were balanced Spanish-English bilinguals, with years of English exposure. At the time of the experiment, though, they were immersed in an English environment. A language frequency effect was found for both languages in the N400 window, as found by Hoshino and Thierry (2012) and Kerkhofs and colleagues (2006), indicating that lexical access occurred as opposed to only language membership information. Additionally, in semantic-category decisions, similar results to Yiu and colleagues (2015) were found when English was the target language, but opposite results when the target language was Spanish. As their participants were, at the time of the experiment, immersed in an L2 environment, the authors argue that language membership is accessed before semantics in the language which is more frequently used in one's certain lifetime period. Even though this issue was not tested directly in our study, our results are in line with Ng and Wicha's (2013). First, if our participants accessed language membership first, there would be no interference of the German reading of the IH, as it would be immediately discarded for being the non-target language of the experiment. As Von Studnitz and Green (2002) explained in their study, this strategy would implicate an internal locus of control and, consequently, a selective language access. In addition, electrophysiological results should show no difference between the control and experimental group, as in the study by Rodriguez-Fornells et al. (2002). Our findings went in the opposite direction: just like Ng and Wicha's (2013) results found in Spanish, the dominant language of their participants, results of our multilingual participants indicated that lexical information was accessed rather than language membership information, since they could not inhibit German's lexical representations, which was their dominant language. In this way, we provide support to BIA+'s prediction that the language nodes are rather represented through an item's language membership "label", which lacks a functional role. Our results provide support to this claim to the extent that, even though our participants may have tried to ignore the words labeled "German" in their mental lexicon, once they were instructed they would read only English words, they suffered an interference from their German semantic information.

4.5 Limitation and future directions

There are some limitations to this study that should be mentioned. Our sample size is relatively small as well as our number of stimuli. However, the type of bilingualism and multilingualism (Portuguese-English/German; late and not in immersion) is not often reported in similar studies. Considering our stimuli list, on the one hand, we strictly controlled some factors which were not controlled in previous studies, such as: a) we had same word class in both readings of the IH and the other stimuli, avoiding nouns which were also verbs in English (this is a very common phenomenon of the language and could not be completely erased); b) we did not admit words with a Latin origin, in order to avoid interference with the native language. This has, on the other hand, unfortunately reduced dramatically the candidate IHs due to the above-mentioned selection criteria, and consequently the number of trials due to both languages' nuisance factors. Another caveat was the difference in familiarity between conditions with and without homographs. Although the effect size was small-medium this factor may have affected RTs and N400 amplitude. This problem may explain why, although our conditions were matched statistically in frequency, our behavioral results did not show a perfect pattern for the control (bilingual) group (e.g.: conditions with IH being slightly harder for both groups).

In spite of the above-mentioned limitations, our study has successfully replicated findings in the bilingual literature for multilinguals, suggesting that lexical access in foreign languages occurs in a language non-selective fashion, with the co-activation of target and non-target representations. This was especially evidenced in the unrelated in English and related in German condition, which indicates that German-speaking participants had some hard time inhibiting German word representations. In line with results of IH in sentence context (ELSTON-GÜTTLER; GUNTER;KOTZ 2005) these results indicate a context-dependent lexical selection phase: there was only disambiguation on the IH selection when its English reading was related to the target. Electrophysiological results show that the presence of IH only affected the group which could "read" them in two competing ways. Further studies should indicate which factors affect IH reading with multilingual speakers.

4.6 Conclusions

In this study, our aim was to investigate the occurrence of co-activation of the dominant foreign language of multilinguals during a task in their non-dominant foreign language, as it has been found in similar studies with monolinguals and bilinguals. The behavioral results showed that the presence of the homographs slowed responses and elevated error rates, especially in the multilingual group. We believe this result is due to difficulties in the inhibition of non-target representations. Moreover, a clear N400 effect only in the bilingual group may indicate that, even though the homographs also interfered with the bilingual group, both groups processed the interlingual homographs differently, which can be taken as evidence of a modulation of the component due to the access of non-target readings in the multilingual group.

Appendix

Characterization of the subjects

All participants were recruited through social media and institutional e-mails for this study. They had normal or corrected to normal visual and auditory acuity, suitable for task demands, and agreed in taking part of the research, by filling in the Written Informed Consent. The socioeconomic status was evaluated by the Brazilian Classification Criteria questionnaire developed by the Brazilian Association of Research Companies (ABEP, 2018), which classifies individuals into seven classes (in descending order of socioeconomic status): A (45-100 points), B1 (38-44), B2 (29-37), C1 (23-28), C2 (17-22) and D-E (1-16) and E (0-7). Most of the participants of both groups belonged to the middle classes B2 and C1: the average score in ABEP Questionnaire was C1 (28.13, SD 24.58) for the bilingual group and B2 (31.07, SD 24.11) for the multilingual group. Regarding participants' educational area, most of them either had a degree or were studying to obtain a college degree in Human Sciences (66.7% of the bilingual participants and 63.6% of the trilingual participants). All participants reported having both parents with Portuguese as their first language.

Foreign Language knowledge and use

Table 1 shows details of age and context of acquisition of English and German and Table 2 represents self-reported proficiency for both groups.

Table 1 – Self-reported language acquisition

	Age of Language Acquisition English N (%)	Age of Language Acquisition German (years) N (%)	Context of Language Acquisition English N (%)	Context of Language Acquisition German N (%)
Bilinguals	1-4 y 2 (16.6%) 5-10 y 3 (25%) 11-20 y 7 (58.3%) 20+ y n/a	1-4 y n/a 5-10 y n/a 11-20 y n/a 20+ y n/a	LI 8 (66.6%) ISch n/a IS 4 (33.3%) Family n/a	LI n/a ISch n/a IS n/a Family n/a
Multilinguals	1-4 y 2 (18.1%) 5-10 y 6 (54.5%) 11-20 y 3 (27.2%) 20+ y n/a	1-4 y 1 (9.0%) 5-10 y 1 (9.0%) 11-20 y 7 (63.6%) 20+ y 2 (18.1%)	LI 8 (72.7%) ISch 1 (9.0%) IS (9.0%) Family 1 (9.0%)	LI 5 (45.4%) ISch 1 (9.0%) IS 3(27.2%) Family 2 (18.1%)

LI: Language Institute (3-5 hours per week); ISch: International School; IS: Individual Studies; y= years; n/a= non-applicable

Source: Own authorship

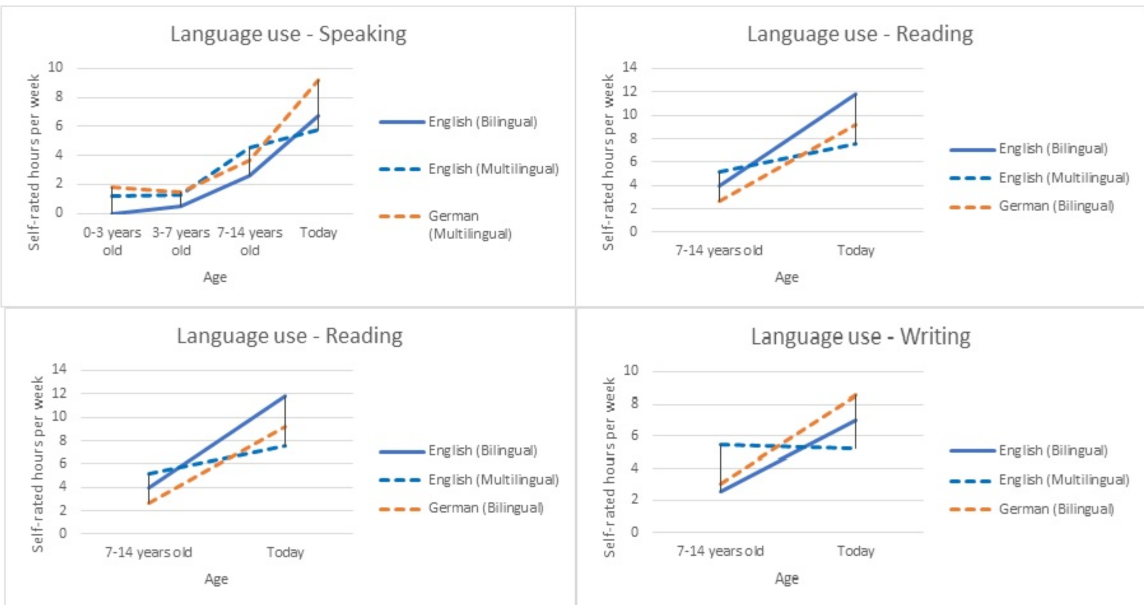
Table 2 – Self-reported language proficiency

Subjective Proficiency English								
	Bilinguals				Multilinguals			
	LST	SPK	RD	WRT	LST	SPK	RD	WRT
Fluent	7 (58.3%)	3 (25.0%)	6 (50.0%)	2 (16.6%)	5 (45.4%)	3 (27.2%)	5 (45.4%)	3 (27.2%)
Advanced	5 (41.6%)	7 (58.3%)	6 (50.0%)	7 (58.3%)	3 (27.2%)	3 (27.2%)	4 (36.3%)	1 (9.0%)
Intermediate	n/a	2 (16.6%)	n/a	2 (16.6%)	3 (27.2%)	5 (45.4%)	2 (18.1%)	7 (63.6%)
Basic	n/a	n/a	n/a	1 (8.3%)	n/a	n/a	n/a	n/a
Subjective Proficiency German								
	LST	SPK	RD	WRT	LST	SPK	RD	WRT
Fluent	n/a	n/a	n/a	n/a	7 (63.6%)	6 (54.5%)	7 (63.6%)	4 (36.3%)
Advanced	n/a	n/a	n/a	n/a	4 (36.3%)	5 (45.4%)	4 (36.3%)	7 (63.6%)

LST: Listening; SPK; Speaking; RD; Reading; WRT: Writing

Source: Own authorship

Figure 7: Self-rated use of languages for participants in the bilingual and multilingual groups, divided into the time-windows 0-3 years old, 3-7 years old, 7-14 years old and at the time they participated in the research.



Source: Own authorship

Language use of both participant groups across their lifespan is represented in Figure 1, in which we show each group's mean of self-reported speaking hours per week. It shows that, between adolescence and adulthood, the participants of the multilingual group started using more German than English. Besides the information collected through self-report, native and foreign language proficiency was also assessed objectively, through the following tests:

I) Semantic verbal fluency – (animals/min) - participants were required to produce, during a minute, as many names of animals as possible. An independent sample t-test was performed on semantic fluency both in Portuguese and in English. No significant difference was found between groups in Portuguese, (t= .23; df=21; p= .81) nor in English (t= 0.25; df=21; p= .80). A paired ANOVA with Tukey post hoc evidenced both groups performed better in Portuguese than English (ps<.01), but no significant difference between English-German languages on multilingual group in this test was found (all ps> .09).

II) Phonemic verbal fluency – participants were asked to produce, during a minute, as many words as possible starting with specific letters. For the English and Portuguese languages, letters were selected according to the American Version of Verbal Fluency Test: A, F and S. In German, the sound /f/ is more often found in nouns which start with the consonant <v>, and this could cause confusion in the participants. Besides that, words starting with A in German are less frequent than in Portuguese and English. Therefore, we adapted this task adopting a subset of letters suggested by the Fluency test of Regensburg (Regensburger Flüssigkeitstest): B, M and S. An independent sample t-test was performed on phonological fluency and showed no significant difference between groups in Portuguese (t= .11; df=21; p= .91) nor in English (t= 1.34; df=21; p= .19). Paired ANOVA Tukey post hoc showed significant difference only in the bilingual group between Portuguese and English versions of this task(p=.006). The multilingual group had similar performance in all the three languages (all ps>.13).

III) Boston Naming test – we used a short version of the Boston Naming Test with 30 pictures instead of 60: 10 high-frequency, 10 medium-frequency and 10 low-frequency. The pictures were to be named in Portuguese and English (bilingual group) and also in German in the multilingual group. An independent sample t-test was performed and showed no significant differences between groups in Portuguese (t= .05; df=21; p= .95) or in English (t= 1.02; df=21; p= .31). – The paired ANOVA with Tukey post hoc evidenced both groups had better performance in Portuguese than English and German languages (all ps<.001). No significant difference was found between English-German in the multilingual group (p= .68).

Neuropsychological characterization

Participants were assessed with Direct and Inverse Digit Span tests (for the assessment of working memory capacity), and we also applied the subtests of Vocabulary and Cubes of WAIS-III (estimate IQ). Independent t-tests reported no significant difference between groups in the raw scores of these tasks (p> 0.72). The neuropsychological assessments were undertaken only in Portuguese. The instructions of the linguistic pre-tests (Boston Naming Test, Verbal Fluency) were given in the tested language. In summary, the bilingual group performed the neuropsychological tests in Portuguese and then the language objective assessment in Portuguese. After that, participants engaged in the task language mode (GROSJEAN, 2010) (English), for the linguistic tests and during EEG preparation and experiment instructions. The multilingual group did the same procedures for the neuropsychological assessments and objective language assessment of Portuguese. After that, participants engaged in German language mode for the linguistic tests in German. Finally, participants engaged in the task language mode (English), for the objective assessment of English proficiency and during EEG preparation and experiment instructions. Only English was used for the rest of the experiment session and participants were told that the rest of the session and the experiment would require English only.

Description of the semantic judgement task (preliminary study and stimuli psycholinguistic features)

Relatedness between prime and target was initially generated with the aid of the software Sketch Engine and Small World of Words (LEUVEN, 2017). They generate, from a typed input word, a ranking of other words which are commonly related to them. As they are based on research developed with native speakers, we implemented a second phase of normalization of the stimuli, with the assessment of semantic relatedness by bilinguals (Portuguese-English) and multilinguals (Portuguese-English-German). One hundred and eight individuals judged the stimuli pairs regarding semantic relatedness, on a 1-5 like rt scale in which 1 means completely unrelated and 5 completely related. Themean and standard deviation of the related word pairs with homograph was 4.13 ± 0.32 and without homograph 4.06 ± 0.25. The mean relatedness rating of the unrelated word pairs with homograph was 1.79 ± 0.47 and without homograph 1.57 ± 0.40). A Kruskal-Wallis test was performed between related (with/without homograph) and unrelated (with/without homograph) word pairs, for a detailed comparison. Stimuli were matched in pairwise comparison test Dwass-Steel-Critchlow-Fligner (p related= .35, p unrelated= .73) regarding the prime-target relatedness.

Stimuli (primes and targets, separately) were also judged according to their familiarity to participants and concreteness. We examined the individual word stimuli performing two Mann-Whitney tests comparing word frequency, word length, familiarity and concreteness of the stimuli in the related vs unrelated conditions and in the conditions with vs without homographs. The comparisons between related and unrelated pairs evidenced that the stimuli were matched in word frequency – mean= 30.26, SD= 32.876 related; mean= 35.06, SD= 35.891 unrelated; word length - mean= 4.59, SD= 0.945 related; mean= 4.69, SD= 1.123 unrelated; familiarity – mean= 4.54, SD= 0.443 related; mean= 4.59, SD= 0.347 unrelated; and concreteness – mean= 4.02, SD= 0.819 related; mean= 4.01, SD= 0.833 unrelated; (p > .48 in all comparisons). In the comparison between pairs with and without homographs, the stimuli were also matched in word frequency (p= .37) – mean= 31.42, SD= 33.923 with homographs; mean= 33.91, SD= 35.002 without homographs. However, in the conditions with homographs the stimuli presented in average fewer letters (Mean= 4.43, SD= 1.03) compared to the conditions without homographs (4.85, SD= 0.99). Although familiarity means in all conditions was very high, the stimuli in the conditions with homographs had a lower familiarity rating (Mean= 4.48, SD= 0.41) than the conditions without homographs (Mean= 4.65, SD= 0.36). Also, conditions with homographs had a higher concreteness rating than without (Mean= 3.53, SD= 0.876 with homograph; Mean= 4.50. SD= 0.363 without homograph). These differences in the homograph conditions were of medium/small effect size (rbp< .25).

References

ABEP. *Critério Brasil*. Available at: <http://www.abep.org/criterio-brasil>.2018.

BAKER, C. *Foundations of Bilingual Education and Bilingualism*, v. 79. Bristol: Multilingual Matters, 2011.

BRYSBAERT, M.; BUCHMEIER, M.; CONRAD, M.; JACOBS, A. M. ; BÖLTE, J.; BÖHL, A. The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, v. 58, n.5, 412-424, 2011. doi:https://doi.org/10.1027/1618-3169/a000123.

BRYSBAERT, M., & NEW, B. Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods*, v.41, n.4, 977-990, 2009. doi: https://doi.org/10.3758/BRM.41.4.977.

DE BRUIJN, E. R. A.; DIJKSTRA, T.; CHWILLA, D.J.; SCHRIEFERS,H.J. Language context effects on interlingual homograph recognition: evidence from event-related potentials and response times in semantic priming. *Bilingualism: Language and Cognition*, v.4, n.02, 2001.doi:10.1017/s1366728901000256.

DIJKSTRA, T. *et al*. More on interlingual homograph recognition: Language intermixing versus explicitness of instruction. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 3, n. 1, 69-78, 2000.

DIJKSTRA, T.; VAN HEUVEN, W. J. B. The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, v.5, n.03, 2002. doi:10.1017/s1366728902003012.

DUDEN. *Duden Online Wörterbuch*. Available at: <https://www.duden.de/woerterbuch>.2017

DURLIK J.; SZEWCZYK J.; MUSZYŃSKI M.; WODNIECKA Z. Interference and Inhibition in Bilingual Language Comprehension: Evidence from Polish-English Interlingual Homographs. *PLoSONE*, v. 11, n. 3, e0151430, 2016. doi:10.1371/journal.pone.0151430.

ELSTON-GÜTTLER, K.E.; GUNTER, T.C.; KOTZ, S.A. Zooming into L2: global language context and adjustment affect processing of interlingual homographs in sentences. *Brain research. Cognitive brain research*, v.25, n. 1, 57-70, 2005.

GROOT, A. M.D.; DELMAAR, P.; LUPKER, S. J. The processing of interlexical homographs in translation recognition and lexical decision: support for non-selective access to bilingual memory. *The Quarterly Journal of Experimental psychology. A, Human experimental psychology*, v.53, n.2, 397-428, 2000. https://doi.org/10.1080/713755891.

GROSJEAN, F. *Bilingual*. [S.l.]: Harvard University Press, 2010.

VAN HEUVEN, W. J.; SCHRIEFERS, H.; DIJKSTRA, T.; HAGOORT, P. Language conflict in the bilingual brain. *Cerebral cortex (New York, N.Y.: 1991)*, v. 18, n. 11, 2706–2716, 2008. https://doi.org/10.1093/cercor/bhn030.

HOSHINO, N.; THIERRY, G. Do Spanish-English Bilinguals have Their Fingers in Two Pies - or is It Their Toes? An Electrophysiological Investigation of Semantic Access in Bilinguals. *Frontiers in psychology*, v. 3, 9, 2012. https://doi.org/10.3389/fpsyg.2012.00009.

HSIEH, M. C.; JEONG, H.; DOS SANTOS KAWATA, K. H.; SASAKI, Y.; LEE, H. C., YOKOYAMA, S.; SUGIURA, M.; KAWASHIMA, R. Neural correlates of bilingual language control during interlingual homograph processing in a logogram writing system. *Brain and language*, v.174, 72–85, 2017. https://doi.org/10.1016/j.bandl.2017.06.006.

KEMMERER, D. *Cognitive neuroscience of language*. [S.l.]: Psychology Press, 2014.

KERKHOFS, R.; DIJKSTRA, T.; CHWILLA, D. J.; DE BRUIJN, E. R. Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects. *Brain research*, v.1068, n.1, 170–183, 2006. https://doi.org/10.1016/j.brainres.2005.10.087.

KOUSAIE, S.; PHILLIPS, N. A. Age-related differences in interlingual priming: a behavioral and electrophysiological investigation. *Neuropsychology, development, and cognition. Section B, Aging, neuropsychology and cognition*, v. 18, n. 1, 22–55, 2011. https://doi.org/10.1080/13825585.2010.510555.

KROLL, J. F.; DUSSIAS, P. The comprehension of words and sentences in two languages. In T. Bhatia & W. Ritchie (Eds.). *Handbook of Bilingualism* Cambridge, MA: Blackwell Publishers, 2004, 169-200.

KUTAS, M.; FEDERMEIER, K. D. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual review of psychology*, v.62, 621–647, 2011. https://doi.org/10.1146/annurev.psych.093008.131123.

KUTAS, M.; PETTEN, C.V.; KLUENDER, R.E. *Psycholinguistics Electrified II (1994–2005)*, 2006.

LEMHÖFER, K.; DIJKSTRA, T. Recognizing cognates and interlingual homographs: effects of code similarity in language-specific and generalized lexical decision. *Memory & cognition*, v. 32, n. 4, 533–550, 2004. https://doi.org/10.3758/bf03195845.

LEMHÖEFER, K.; DIJKSTRA, T.; MICHEL, M. Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, v. 19, n. 5, 585-611, 2004. https://doi.org/10.1080/01690960444000007.

LEUVEN, B. University of. *Small World of Words*. Available at<http://www.smallworldofwords.com/new/visualize/>. 2017.

MACIZO, P.; BAJO, T.; MARTÍN, M.C. Inhibitory processes in bilingual language comprehension: Evidence from Spanish-English interlexical homographs. *Journal of Memory and Language*, v. 63, 232-244, 2010.

MACMILLAN. *Macmillan Online Dictionary*. Available at<http://www.macmillandictionary.com/dictionary/american/>. 2017.

MARTÍN, M.C.; MACIZO, P.; BAJO, T. Time course of inhibitory processes in bilingual language processing. *British journal of psychology, 101 Pt 4*, 679-93, 2010.

NG, S.; WICHA, N. Y. Meaning first: a case for language-independent access to word meaning in the bilingual brain. *Neuropsychologia*, v. 51, n. 5, 850–863, 2013. https://doi.org/10.1016/j.neuropsychologia.2013.01.017.

RODRIGUEZ-FORNELLS, A.; ROTTE, M.; HEINZE, H. J.; NÖSSELT, T.; MÜNTE, T. F. Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*,v. 415, n. 6875, 1026–1029, 2002. https://doi.org/10.1038/4151026a.

STUDNITZ, R.E.; GREEN, D.W. Interlingual homograph interference in German–English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition*, v. 5, 1 – 23, 2002.

THIERRY, G.; WU, Y. J. Electrophysiological evidence for language interference in late bilinguals. *Neuroreport*, v. 15, n. 10, 1555–1558, 2004. https://doi.org/10.1097/01.wnr.0000134214.57469.c2.

THIERRY, G.; WU, Y. J. Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences of the United States of America*, v. 104, n. 30, 12530–12535, 2007. https://doi.org/10.1073/pnas.0609927104.

TOASSI, P. F. P.; MOTA, M. B.; TEIXEIRA, E. N. The effect of cognate words on lexical access of English as a third language. Cadernos de Tradução [online], v. 40, n. 2, 74-96, 2020. Available from: <https://doi.org/10.5007/2175-7968.2020v40nesp2p74>.

WISSENSCHAFTEN, B.-B. A. der. *DWDS Online Wörterbuch*. Available from: <http://www.dwds.de/>. 2017.

WU, Y. J.; THIERRY, G. Investigating bilingual processing: the neglected role of language processing contexts. *Frontiers in psychology*, v. 1, 178, 2010. https://doi.org/10.3389/fpsyg.2010.00178.

YIU, L. K.; PITTS, M. A.; CANSECO-GONZALEZ, E. Electrophysiological assessment of the time course of bilingual visual word recognition: Early access to language membership. *Neuropsychologia*, v. 75, 349–367, 2015. https://doi.org/10.1016/j.neuropsychologia.2015.06.018.

ZHU, Y.; MOK, P. P. K. Visual recognition of cognates and interlingual homographs in two non-native languages. *Linguistic Approaches to Bilingualism*, 2018. doi:10.1075/lab.17049.

Recebido em 13 de setembro de 2022
Aceito em 29 de dezembro de 2022