PETRA IHÁSZ[1] – ANDRÁS BENYHE[2] – GYULA SÁRY[2] –
ZOLTÁN JUHÁSZ[1] – JUDIT NAVRACSICS[1]

[1] University of Pannonia
[2] University of Szeged
Correspondence author: Ihász Petra, ihasz.petra@mftk.uni-pannon.hu
ORCID-id: 0000-0003-3515-8230

Petra Ihász – András Benyhe – Gyula Sáry – Zoltán Juhász – Judit Navracsics: Visual Word
Recognition Patterns of Hungarian-English Bilinguals – Homograph Effect
in Bilingual Language Decision
Alkalmazott Nyelvtudomány, XXIII. évfolyam, 2023/1. szám, 36–58.
doi:http://dx.doi.org/10.18460/ANY.2023.1.003

Visual Word Recognition Patterns of Hungarian-English Bilinguals – Homograph Effect in Bilingual Language Decision[1]

The present study is part of a larger-scale research in which the processes of written word recognition are studied in bilinguals. The research goal of our lexical decision experiments is to gain information about the temporal characteristics of recognition at the orthographic, phonological, and semantic levels of processing. The research questions concern behavioral differences and the ERP components of recognizing English words, Hungarian words, and interlexical homographs. 23 Hungarian-English bilingual participants were tested in an Electroencephalogram laboratory. In recognition of Hungarian and English words and homographs, the mean response language per participant indicated high accuracy for both Hungarian and English conditions (96% and 98%, respectively). In contrast, the homographs are biased towards English responses (27% Hungarian response). The multiple comparisons confirmed no difference in the mean response times of Hungarian and English words, whereas the interlexical homographs produced around 150 ms longer responses. In recognition of Hungarian and English words, there was no difference between the two categories in the early recognition phases, corresponding with the orthographic-phonological level. However, the neural representation of the two languages differed, later reflecting the differences in semantic or decision-related processes. In the case of the Hungarian-English interlexical homographs, the ERP waveforms did not show significant differences between the items perceived as English or Hungarian. Although there is a difference between the brain activations in the temporal and frontal electrode sites, this difference is insignificant. These data coincide with the former findings related to the homograph effect (Navracsics & Sáry, 2013), which explains that participants are exposed to a greater cognitive burden in the recognition, and the reaction time is longer due to the fact that both lexicons are active.

Keywords: EEG, ERP, bilingual visual word recognition, interlexical homographs, language decision test

1. Introduction

1.1 Visual word recognition and reading

Bilingual written language processing can be studied by observing the brain activations during the processing of the two languages in a bilingual mode, paying particular attention to the frontal, temporal, parietal, occipital, occipito-parietal lobes, as well as the central parts of the brain.

---

[1] We acknowledge the financial support of Széchenyi 2020 under the EFOP-3.6.1-16-2016-00015.

The visual cortex, located at the back part of the brain, is the first brain area where the processing of visual information starts. The ventral pathway, which includes several cortical and subcortical areas, also has a significant role in visual recognition. All these areas create neural activities in the visual processing areas, which process different aspects of perception, such as shape, color, depth, location, movement, etc. (De Groot, 2011). These pieces of information then go to the temporal lobe, where the recognition occurs. Later, some information travels to the frontal lobe, revealing its significance and meaning. At this point, different components connect, and at the end of the whole process, a conscious recognition occurs (Yamins et al., 2014).

The initial reading stage also takes place in the visual cortex, which sends the information toward the brain's language areas. The information arrives at the word-recognition area, which can distinguish between objects and written words. In the auditory cortex, written words are transformed into phonological elements so they can be ‘heard’ inside. Broca’s area is the center of recognizing written words as meaningful utterances by connecting written and spoken words. The information arrives at the temporal lobe, which matches the words to their meanings by retrieving memories (Carter, 2009).

The hierarchy of the language system influences reading. The visual form of words (orthography) is the most influencing factor in word recognition, even for an experienced reader (Csépe, 2006). Besides orthography, phonology and semantics also play an essential role in visual word recognition.

Visual language processing is exceptionally significant for bilinguals learning to read since written language comprehension is based on successful word recognition. Although identifying written words is well-researched in monolingual contexts, bilingual visual language processing is still under-researched, especially with Hungarian as a component of bilingualism. Research into bilingual word processing can provide crucial information for researchers and teachers who deal with bi- or multilingual children and facilitate their literacy development. The perception of the visual word is a fundamental skill in everyday activities such as reading. Studies concerning visual word recognition of bilinguals are essential since numerous bilingual students attend monolingual educational institutions, and teachers have to be aware of the processes in the bilingual students’ minds while reading since they have to cope with two languages. Word recognition patterns of orthographically related languages (e.g., English and Dutch) are presumably the same at lower (orthographic and phonological) levels. However, semantic recognition is strongly language-specific. In the case of orthographically unrelated languages (e.g., Hungarian and Chinese), language-specific characters help the recognition process. Hungarian has a shallow writing system; there is a grapheme-phoneme correspondence in it, unlike English, which has a deep one. The present study aims to shed light on the temporal aspects of Hungarian and English written word recognition in the bilingual mind.

1.2 Storage hypothesis and the bilingual mental lexicon

The mental lexicon does not contain words in alphabetical order (Aitchison, 1987). Instead, information is stored in a thematic arrangement. The mental lexicon contains all the information (phonological, morphological, syntactic, semantic) about individual words or morphemes. The bilingual mental lexicon is the key to understanding the nature of language organization in the brain. For a considerable time, the question of the storage of languages has been in focus in the psycholinguistics aspects of bilingualism research (Navracsics, 2007; Pavlenko, 2009; Singleton, 1999). The study of the bilingual mental lexicon and how we store the languages are closely related to brain lateralization.

Early aphasics research showed equal involvement of the two hemispheres in language production and perception (Zatorre, 1989; Solin, 1989). Recent neuroimaging findings deny it, and researchers claim that there are significant individual differences in storage and processing. The relationship between the two languages of a bilingual also varies from individual to individual as the age and manner of second language acquisition and language proficiency level are crucial influencing factors.

The five most common hypotheses on brain lateralization in the case of bilinguals (Hull & Vaid, 2005; Vaid & Hall, 1991) are the following: (i) L2 hypothesis: the right hemisphere is more involved when bilinguals process their L2 than when they process L1. In processing L1, their left hemisphere is involved to the same extent as in language processing by monolinguals. (ii) Balanced bilingual hypothesis: in both L1 and L2 processing, high-proficient bilinguals use their right hemisphere more than monolinguals. (iii) Stage of L2 acquisition hypothesis: in the initial stages of L2 acquisition, the right hemisphere is more involved in processing, and the involvement of the left hemisphere grows with the increase of L2 proficiency (Obler, 1981). (iv) Manner of L2 acquisition hypothesis: if the bilingual individual informally acquires L2, the right hemisphere is more involved than in a formal manner. (v) Age of L2 acquisition hypothesis: if there is only a small gap between the acquisition of L1 and L2, the lateralization pattern will be similar for both languages, i.e., early bilinguals show a similar, while late bilinguals – a different pattern for their two languages (Vaid & Genesee, 1980).

De Groot (2011) assumes that the left hemisphere plays a more prominent role in linguistic behavior, so language is lateralized in the left hemisphere in most people. However, there are other presumptions, according to which manner and age of acquisition play an important role in hemisphere lateralization. Suppose the second language is acquired in childhood. In that case, it is more semantics-based, and the left hemisphere is more involved, while the more acoustics-based second language learning in adulthood results in the right hemisphere being more involved in production and perception (Marrero et al., 2002). Informal second language acquisition involves the subcortical structures, such as the basal ganglia and the cerebellum, and the two languages have shared storage, while second language learning through an instructional way is located in the cerebral cortex; hence L1 and L2 are stored separately (Fabbro & Paradis, 1995; Fabbro, 2000).

In conclusion, the later the second language acquisition starts, the greater the difference between the two hemispheres' lateral organizations. Mechelli et al. (2004) reveal that bilingual adults have greater gray matter density, especially in the inferior frontal cortex of the brain’s left hemisphere, which is the center of language and communication. This type of increased density is observable in bilinguals who started learning their second language before the age of five. Hull and Vaid (2007) support this idea in their meta-analysis of 66 healthy subjects. They discovered that the age of acquisition determines functional lateralization.

Both structural and functional imaging studies (McLaughlin et al., 2004; 2010) show that the brains of adult L2 learners change before their behavior actually realizes the learning processes, and they confirm that these changes are dynamic over time. Furthermore, recent neuroscience evidence (Bice & Kroll, 2015; Chang, 2012, 2013) indicates that L2 begins to change L1, even in the beginner stage of L2 learning. Ameel et al. (2005) find that L1 does not look strictly the same for bilinguals as for monolingual speakers of the very same language. Co-activation of both languages occurs at all levels of language processing, such as lexicon (Malt et al., 2015), grammar (Dussias & Scaltz, 2008), and phonology (Goldrick et al., 2014).

1.3 Visual word recognition models

The main focus of bilingual visual word recognition is the neurocognition of multiple languages. The main questions are the following: (i) how bilinguals select between their languages; (ii) how the conceptual meanings are associated with individual words; and (iii) whether each language has a separate conceptual storage space or not.

1.3.1 The Bilingual Interactive Activation+ (BIA+) model

Several models demonstrate bilingual word recognition, such as the Bilingual Interactive Activation (BIA) model and the Semantic, Orthographic, and Phonological Interactive Activation (SOPHIA) model. Still, according to our research, the Bilingual Interactive Activation+ (BIA+) model is the most relevant. In this model, the input is checked by the sublexical orthography and sublexical phonology, which continuously interact with each other. Then the information is forwarded to the subsystem of lexical orthography and lexical phonology, which are also connected. The language nodes and semantics check the input, and the language gets selected. It is an interactive model since all levels are interconnected, and there is transparency between the subsystems. The information can be sent back to the previous subsystem to confirm. When the target language is chosen, the semantics of the word is checked. The task/decision subsystem receives the information from the word identification subsystem, where the correct language is identified and activated (Dijkstra & Van Heuven, 2002).

1.4 Single-word processing

An efficient method for uncovering the neural background of word recognition is electroencephalography (EEG) and the extraction of Event-Related Potential (ERP) components. EEG is a non-invasive technique which measures brain electrical activity with high temporal precision and a limited but useful spatial localization. By averaging together the EEG data in a stimulus-onset-centered timeframe, we can assess the changes in electrical activity related to the processing of a stimulus (the ERP waveform). After the onset of a visual stimulus, the visual system gets activated, indicated by a series of positive and negative deflections in the ERP waveforms in the occipital electrode sites. The first such wave is the P100 component (positive deflection at 100 ms), related to lower-level visual processing (primitive shapes, possibly letters). The following crucial component is the N170 (negative deflection peaking around 170 ms) that reflects higher-level processing to the visual processing of words. The N400 (negative deflection peaking at 400 ms) is a more centrally located ERP component that reflects the semantic processing of words and other meaningful stimuli. This moment is when lexico-semantic processing identification occurs, with less expected words producing higher amplitude N400 (Carreiras et al., 2013).

In language decision tasks, pseudo-words evoke larger amplitude N400s than words (Braun et al., 2006). According to Braun et al. (2006), the amount of neural activity depends on two critical factors: (i) the difficulty of the visual word processing itself; (ii) neural activity is affected by the global amount of information.

1.5 The recognition of interlexical homographs

Orthographically identical but phonologically and semantically different words in the two languages are interlexical homographs (e.g., *comb, eleven,* etc. in English and Hungarian). A particular subcategory of interlexical homographs are cognates (e.g., *film, café, farm, park, opera, taxi*, etc.), which have identical spelling and share meanings across languages (De Groot, 2011).

Psycholinguistic studies of bilingual language processing agree that representations from different languages (having alphabetic orthographical system) are simultaneously activated, and bilinguals cannot completely deactivate either of their languages, and the information in the other language is also being assessed (Kroll et al., 2015; Van Heuven & Dijkstra, 2010). Previous findings have confirmed that cross-language interaction exists in the case of bilinguals while reading, listening, and speaking, regardless of their proficiency levels (Kroll & De Groot, 2005). ERP studies have also proved that there is a parallel activation of lexical information of the two languages (Elston-Guttler et al., 2005), especially in the case of interlexical homographs, since they have unique cross-linguistic features (Studnitz & Green, 2002).

The general purpose of presenting homographs is to discover if lexical activation is embedded in the language (language-selective) or not (language-nonselective). To be more specific, the question is whether both meanings are activated or only the contextually appropriate language when an interlexical homograph is presented to a bilingual.

Beauvillain and Grainger (1987) were the first to test how bilinguals process interlexical homographs in isolation. They used a cross-language primed lexical decision test, in which a set of stimulus pairs was presented to English-French bilinguals. The stimulus pairs contained a French prime word, and an English target word (or non-word), and the words were presented successively. The participants were asked to read each prime and then make a lexical decision on the following target. Most primes were French words, but some were English-French interlexical homographs. The researchers found that in the beginning, both meanings of the interlexical homograph primes were activated, and after a little while, the inappropriate meaning was deactivated. Both lexicons were activated since bilinguals participated in the task in a bilingual processing mode. Green and Abutalebi (2013) introduced the Adaptive Control Hypothesis, which says that the degree of activation is dynamically adaptive. The hypothesis relies on the fact that the state of bilingual mode alters according to the context. This was also confirmed by Grosjean (1998, 2001), whose Language Mode model indicates that bilinguals experience different states of activation of their languages and language processing mechanisms at a given time. De Groot (2011) presents the Language Mode theory as an explanation for the language-nonselective processing of interlexical homographs.

Researchers also tried to find proof for co-activation in the non-target lexicon without suspecting the dual-meaning activation theory. In the study of Kerkhofs et al. (2006), responses to interlexical homographs and unilingual control words (words existing only in the target language) were compared with each other. Features that might influence word processing were monitored, and they found that word frequency is a salient contributing factor.

The lexical decision task is one of the most frequently used methods of testing bilingual visual word recognition. It reveals differences in reaction time and the number of errors between interlexical homographs and control words (Navracsics & Sáry, 2013). In the Hungarian-English bilingual visual word recognition study of Navracsics and Sáry (2013), the homograph effect was observable: the reaction time of the recognition of homographs was significantly longer than that of non-homographs. The researchers also found that the reaction time increased when participants recognized them as Hungarian words (0.94-1.04s). In contrast, recognizing homographs as English words took shorter (0.86s), though the difference was insignificant. They concluded that the increased reaction time in recognition of homographs is because more semantic areas are involved. They also discovered that decision-making in the case of homographs highly depended on word frequency, similar to the findings by Dijkstra et al. (2000) and De Groot (2011). The reaction time in recognition of homographs is longer if their meaning is more frequent in the non-target language. The rejection of the non-target meaning and the access to the appropriate language increase recognition time.

1.6 The present study

The present study seeks to investigate the temporal characteristics of written word recognition. The research questions concern the behavioral and electrophysiological correlates of recognizing English words, Hungarian words, and interlexical homographs.

For this study, the following research questions were formulated: (i) do highly proficient L2 users have the same latencies of language recognition in both their languages; (ii) do the neural responses differ during the recognition of L1 and L2 words; (iii) does the language recognition of homographs differ from non-homographs; (iv) what are the influencing factors of language recognition.

Our hypotheses are as follows: (i) highly proficient L2 users have the same latencies of word recognition in both their languages; (ii) the recognition of the two languages has the same activation patterns; (iii) the recognition of homographs is longer than non-homographs; (iv) word frequency and language-specific characters have a decisive role in visual word recognition.

2. Methods

2.1 Participants

Twenty-three Hungarian–English bilingual volunteers (10 males, mean age: 24.57 yrs, 19 right-handed) were tested in an EEG laboratory. All of them were native speakers of Hungarian with C1 level English proficiency, and they all use English at work and in their everyday lives. They spend at least half an hour a day reading English books and articles. The majority actively uses English for several hours a day on average. None of the participants lived in an English-speaking country for over three months. They come from Hungarian monolingual families and use Hungarian at home. All of them are late bilinguals; they started acquiring English in an instructed manner at primary or secondary school (the mean age of acquisition is 9.97 years). They all had normal or corrected-to-normal (glasses or contact lenses) vision; no hearing impairment, language disability, learning disability, or any history of neurological illness was reported.

2.2 Procedure

The study had the approval of the Local Ethics Committee. Before the test, a consent form and instructions were handed to each participant, who had to read and sign it. Participants were informed that the experiment takes approximately one hour, it is non-invasive, i.e., it does not cause physical pain or inconvenience, and they can interrupt the experiment at any time without any consequences.

Before the EEG experiment, participants completed a non-standardized language background questionnaire about their Hungarian and English language use. They also completed a standardized questionnaire (Language Experience and Proficiency Questionnaire – LEAP-Q), in which they had to list all the languages they know in order of dominance, list all the languages they know in order of acquisition, list the percentage of time they currently and on average are exposed to each language, whether they have lived abroad for a more extended period, etc. (Marian et al., 2007).

Participants were asked to minimize any movement during the test to diminish the data's noisiness.

2.3 Test materials

The language decision test included 180 monosyllabic words: 60 Hungarian (e.g., *bál, cím, lyuk*), 60 English (e.g., *age, cat, hair*), and 60 interlexical homographs (e.g., *comb, hold, mind*) and cognates (e.g., *blog, film, lift*). We used the Hungarian National Corpus (HNC) for Hungarian and the Corpus of Contemporary American English (COCA) for English to control for word frequency. We calculated the Zipf-frequencies of all items as the ten-base logarithm of the frequency per billion words. The Zipf-frequency of Hungarian words was 4.29 (±0.76 SD) and that of English – 4.77 (±0.42 SD) in their respective corpora. The Zipf-frequency of homographs was 4.25 (±0.88 SD) in the Hungarian corpus and 4.6 (±0.80 SD) in the English corpus, and the Hungarian-English frequency difference was -0.35 (±1.00 SD). Since all participants were Hungarian, they were familiar with all Hungarian words. According to the Oxford Dictionary (www.oxforddictionaries.com), all English words in the list belong to A1-B1 levels. Participants were asked to decide whether the word on the screen was Hungarian or English and click on the computer mouse's left (English word) or right (Hungarian word) button. Words appeared on the screen in a mixed, pseudorandom order to keep participants’ both languages active.

2.4 Stimulus presentation and EEG recording

A previously designed custom-made program written in MATLAB (MatLab Inc.) with the Psychtoolbox extension (Kleiner et al., 2007) running on a PC was used for the experiments (Navracsics and Sáry, 2013). Stimuli were presented on a white background, using black characters in the middle of the screen. The viewing distance was set to be the appropriate average viewing distance of a computer screen (~ 50 cm). Trials started with the onset of a fixation spot in the middle of the screen, which was followed by a stimulus chosen from the pool. The inter-trial interval was 1 s; the stimulus stayed on the screen for 2 s (exposure time). During this time, participants were requested to click the right or left button according to the task instructions. Failure to respond in the time window resulted in the continuation of the task to the next trial. The task was machine paced to ensure the participants' constant level of attention. The program recorded the response side (language) and response latency times. Neural activity was recorded with a 128-channel EEG system (Biosemi).

2.5 Data analysis

Incorrect responses were excluded from the analyses (note: for the homographs, all responses were regarded as correct since they can be understood in both languages). Response times and response languages were averaged separately per condition (Hungarian, English, and homograph) for each participant. Language bias of homographs was tested by comparing the rate of Hungarian responses to 50% with Student’s T-test. The mean response times were compared among conditions with repeated measures ANOVA, and post hoc testing was performed with multiple comparisons.

The ratio of Hungarian responses for the homographs was calculated across participants. This item-wise mean language response was tested for linear correlation (Pearson) with the items’ difference between English and Hungarian Zipf-frequencies.

The response times of homograph trials were further divided into two groups based on the language decision and averaged per participant. The means were compared with a paired Student’s T-test. The linear relationship between response time bias (response time difference between Hungarian and English responses to homographs) and decision bias (the ratio of Hungarian responses to homographs) was assessed by calculating the Pearsons correlation coefficient.

The EEG data were preprocessed by re-referencing to the average of all channels, removing line noise with a band-stop filter around 50Hz and band-pass filtering with a 0.5-30 Hz FIR filter. Eye movement artifacts were removed manually, observing and excluding noisy ICA components. Stimulus-locked epochs were extracted from -1 s to 2 s around stimulus onset time. Epochs were baselined to the mean amplitude in the -200-0 ms pre-stimulus window and finally averaged in each channel to obtain ERP waveforms.

Data from each participant was processed individually, and group-level analysis took place with the FieldTrip toolbox in MATLAB. The data were compared between the critical conditions (Hungarian vs. English words; homographs with Hungarian vs. English responses). We used a dependent samples T-test with permutation-based cluster correction (1000 Monte-Carlo permutations) across all channels in the 100-600 ms time window to identify significant differences in the grand averaged ERP waveforms. This correction method analyzes data points in the context of their neighbors in the time and location dimensions. Clusters of significant t-statistic ($p < 0.05$) were considered genuinely significant if the cluster size exceeded 97.5% of the randomly permuted cluster sizes.

To compare the N400 component amplitudes, we averaged voltage levels in the time window between 380 and 420 ms post-stimulus onset at the D14 electrode (roughly corresponding to C1 in a 10-10 system). These amplitude values were then averaged by condition (Hungarian, English, and homograph) for each participant. Condition effects were evaluated by repeated measures ANOVA and multiple comparisons, similar to the response time analyses above.

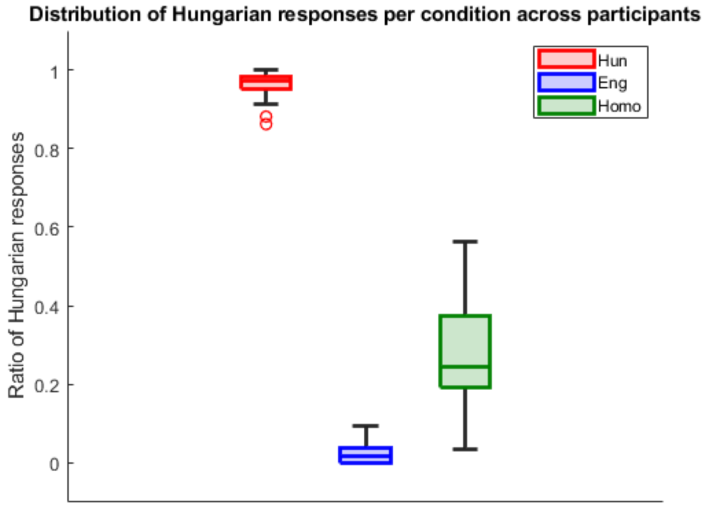
3. Results

The experimental data of one participant was lost, and the following results thus include data from the remaining 22 participants.

3.1 Behavioral measures

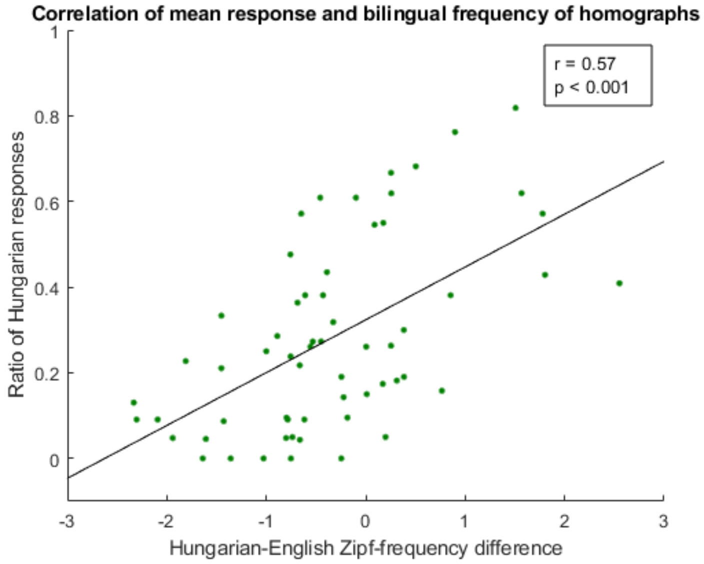
The mean response language per participant indicated high accuracy for both Hungarian (96% correct) and English (98% correct) conditions, whereas the homographs indicated a bias towards English responses (29% Hungarian response; $t(21) = -7.21$, $p < 0.001$) despite the balanced homograph frequencies between the two languages (Fig. 1.).

Figure 1. Distribution of Hungarian response ratios averaged by participant. The boxes display the median, lower, and upper quartiles, and the whiskers extend to the non-outlier minima and maxima. Outliers are defined as data points at least 1.5 inter-quartile range from the top or bottom of the boxes.



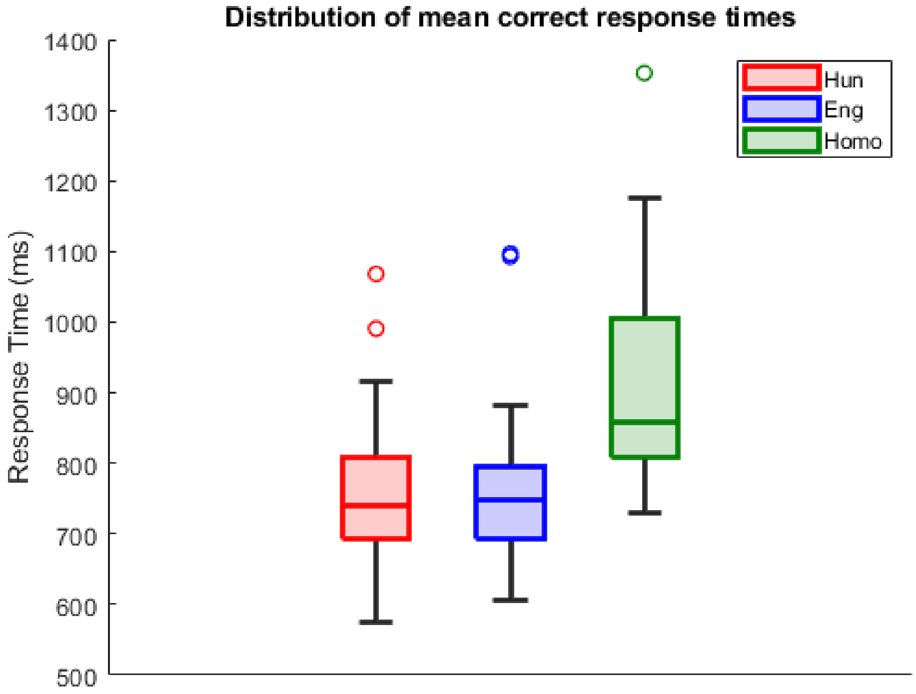
We assessed the relationship between each homograph word's mean response language and the relative frequency with a Pearson's test. The coefficient showed a strong correlation between the ratio of Hungarian responses and the Hungarian-English Zipf-frequency difference (Fig. 2.; r(59) = 0.57, p < 0.001).

Figure 2. Linear correlation between relative frequency and the ratio of Hungarian responses for each homograph item averaged across participants. The fitted line has an intercept of 0.32 and a slope of 0.12. Note that the frequency difference is logarithmic; thus, a value of -1 means that the item is 10 times more frequent in English than in Hungarian, and a value of 2 means that the item is 100 times more frequent in Hungarian than in English.



The mean correct response times were 768 ms, 772 ms, and 922 ms for the Hungarian, English, and homograph conditions, respectively (Fig. 3.). The ANOVA yielded a significant effect of language condition (F(2,21) = 52.59, p < 0.001). The multiple comparisons confirmed that there was no difference in the mean response times of Hungarian and English words (p = 0.94, CI = [-34.39, 26.13]), whereas the homographs produced around 150 ms longer responses than the unambiguous words (Hungarian-homograph: p < 0.001, CI = [-211.30, -98.47]; English-homograph: p < 0.001, CI = [-190.09, -111.42]).

Figure 3. Distribution of correct response times averaged by participants. The boxes display the median, lower, and upper quartiles, and the whiskers extend to the non-outlier minima and maxima. Outliers are defined as data points at least 1.5 inter-quartile range from the top or bottom of the boxes.



The comparison of homograph response times based on language decision revealed a difference between Hungarian and English responses (Fig. 4.). Hungarian responses took, on average, 995 ms. In contrast, for English, they took 916 ms, a difference that proved to be significant upon analysis (t(20) = 3.85, p < 0.001). One participant was excluded from these calculations due to having an extremely low number of Hungarian responses (2 out of 60). The Pearson test revealed a robust linear correlation between the language bias and response time bias of the participants (Fig. 5.; r(20) = -0.84, p < 0.001). This result shows that the less a participant responds to homographs as Hungarian, the slower the Hungarian responses get.

Figure 4. Distribution of homograph response times averaged by participant, based on decision language. The boxes display the median, lower, and upper quartiles, and the whiskers extend to the non-outlier minima and maxima. Outliers are defined as data points at least 1.5 inter-quartile range from the top or bottom of the boxes.

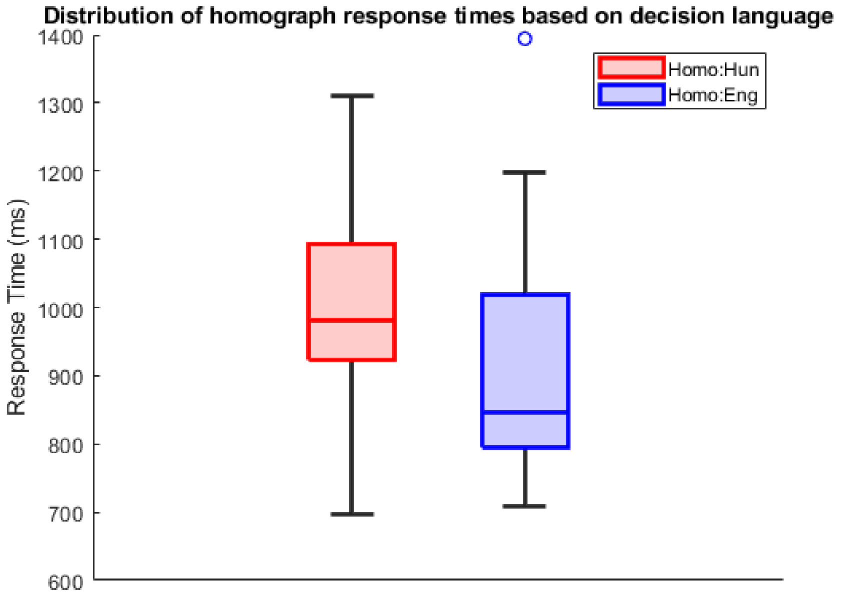
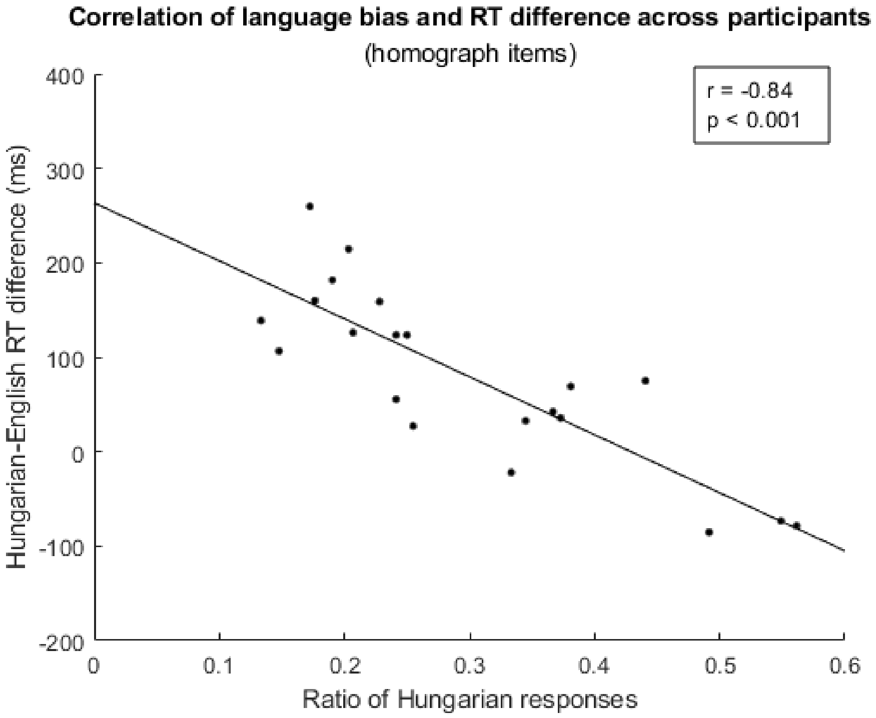


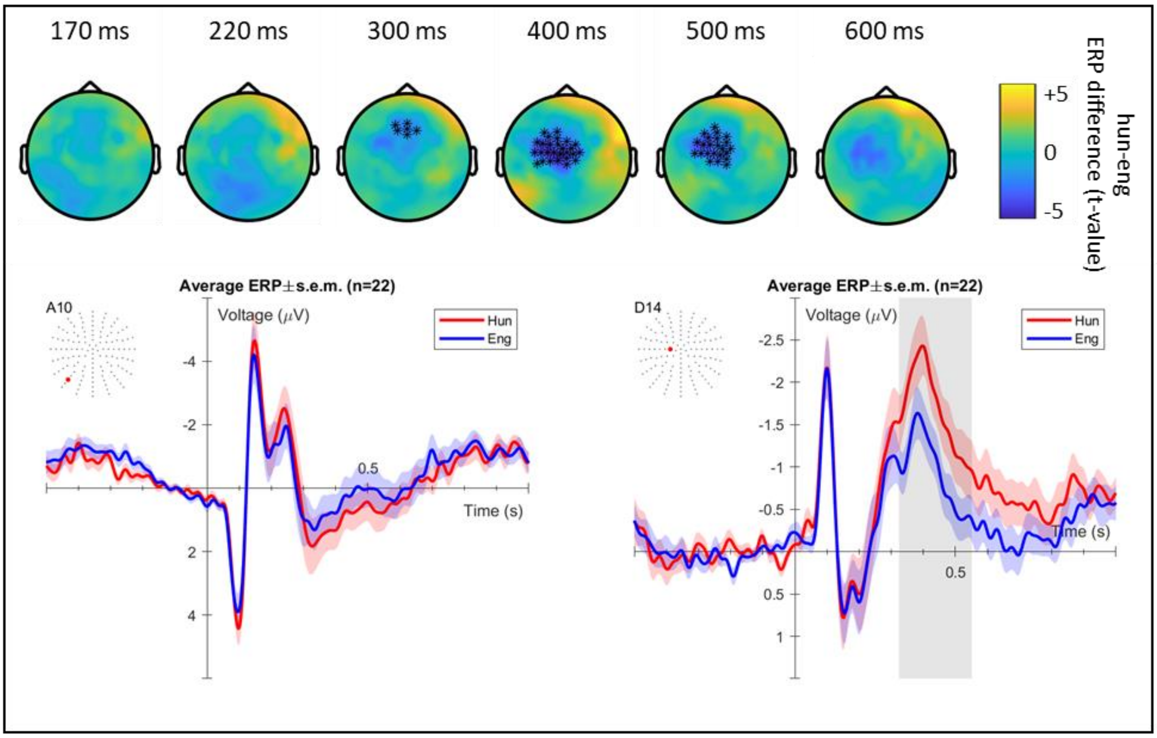
Figure 5. Linear correlation of language bias and response time difference of homographs. The fitted line has its intercept at 263 ms, and the slope is -614 ms.



3.2 ERPs of non-homographs

The ERP waveforms elicited by Hungarian and English words did not seem to differ in the early stages of visual word recognition. The occipital P100 and N170 components are clearly identifiable in the occipital regions (Fig. 6. bottom left), and the cluster-based statistics indicate no differences in this time window between the two conditions. However, the central electrode sites show a difference in the N400 component (Fig. 6. bottom right), with the Hungarian words producing larger (more negative) amplitude. This difference belongs to a significant cluster, spanning from 300 ms to 500 ms (Fig. 6. top).

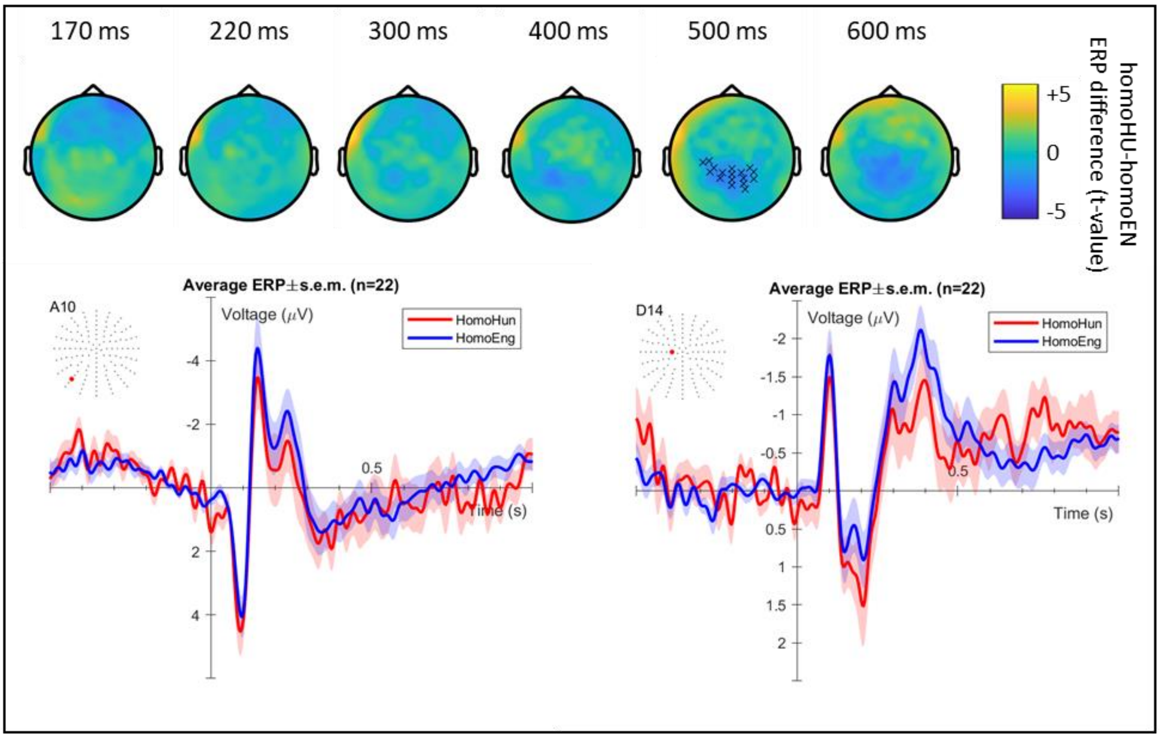
Figure 6. (Top) Topoplots representing the ERP difference between Hungarian and English at denoted times. Channels with significant contrast are denoted by asterisks ($p < 0.01$). (Bottom) ERP waveforms at the left occipital A10 (left panel) and the central D14 (right panel) channels. The shading represents times of significant difference ($p < 0.05$).



3.3 ERPs of homographs

The above-seen N400 difference could not be reproduced with homographs recognized as Hungarian or English, although a weak centro-parietal cluster emerged around 500 ms after stimulus onset (Fig. 7. top). The occipital and central ERP waveforms were not found to differ at any time points (Fig. 7. bottom).

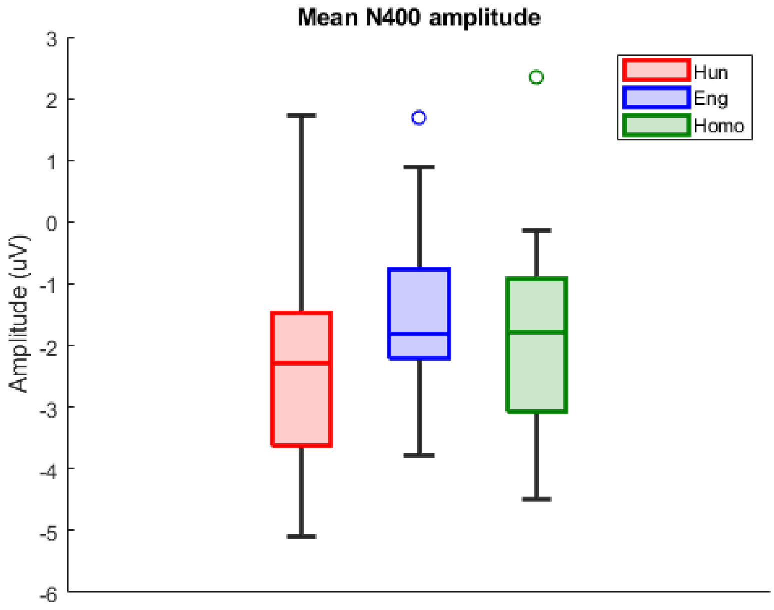
Figure 7. (Top) Topoplots representing the ERP difference between Hungarian-regarded and English-regarded homographs at denoted times. Channels with marginally significant contrast are denoted by crosses ($p < 0.05$). (Bottom) ERP waveforms at the left occipital A10 (left panel) and the central D14 (right panel) channels.



3.4 N400 components

The comparison of the mean N400 components revealed a significant effect of language condition (Fig. 8.; $F(2,21) = 7.79$, $p = 0.001$). The mean component amplitudes were -2.34 uV for Hungarian words, -1.49 uV for English words, and -1.78 uV for homographs. The only significant contrast upon multiple comparisons was seen between Hungarian and English non-homographs ($p < 0.001$, CI = [-1.33, -0.39]).

Figure 8. Distribution of mean N400 component amplitudes averaged by participants. The boxes display the median, lower, and upper quartiles, and the whiskers extend to the non-outlier minima and maxima. Outliers are defined as data points at least 1.5 inter-quartile range from the top or bottom of the boxes.



4. Discussion

The present study investigated a language decision task's behavioral and electrophysiological correlates with Hungarian, English, and interlexical homograph words.

Word recognition patterns of orthographically related languages (e.g., English and Dutch) are presumably the same on lower levels (orthographic and phonological). Still, at higher cognitive levels, recognition is strongly language-specific in semantics. In orthographically unrelated languages (e.g., Hungarian and Chinese), language-specific characters help the recognition process with the language decision. The two languages investigated in this paper have the Latin alphabet. The majority of letters are identical, but there are some language-specific letters with diacritics in Hungarian, making it easy to recognize Hungarian words at the orthographic level.

While Hungarian has a shallow writing system based on a consistent mapping of graphemes to phonemes, English has a deep writing system that lacks the grapheme-phoneme correspondence rule. Hungarian and English are not connected typologically. In the case of bilinguals who speak two typologically dissimilar languages, the language-specific letter string activates the appropriate language instantly since the other language lacks that combination of letters (Singleton, 1999). The identification of the two languages had similar activation patterns in this investigation of highly skilled bilinguals. Our findings are consistent with those of other researchers who have investigated typologically related languages such as Spanish-English (Macizo et al., 2010; Schwartz et al., 2007) or Dutch-English (Lemhöfer & Dijkstra, 2004; Van Assche et al., 2009).

Results also show that word recognition activates different parts of the brain from stimulus onset to word identification, confirming hypotheses about the neurolinguistic and temporal characteristics of bilingual visual word recognition (Navracsics & Sáry, 2013; Carreiras et al., 2013; De Groot, 2011). There is no significant difference between Hungarian and English word recognition on the orthographic-phonological level. It means that participants did not have to put in any extra effort to identify the words, implying that word familiarity is important in visual word recognition, as claimed by Assadollahi and Pulvermuller (2003), Dambacher et al. (2006), and Yum and Law (2021).

We found that the responses to unambiguous words were equally fast and accurate for Hungarian (L1) and English (L2) items. However, the responses slowed drastically (~150 ms) for homograph words and showed a bias towards English responses, even though, on average, the homograph items were equally frequent in both languages. Although the variation in the response language can be partly explained by the relative frequency between the two languages, the skewed nature of the homograph responses is clear, showing a bias towards English.

The reaction times for homograph items were found to be slower for Hungarian responses, in line with the findings of Navracsics and Sáry (2013). This result seems to agree with the previously mentioned response bias, an advantage of English over Hungarian. The two effects line up nicely, with a robust correlation between the language decision preferences and the time cost of Hungarian responses. We propose that this bias is indicative of the underlying strategy that participants developed during the experiment. The task was likely reformulated in many (at least those with a more substantial bias) to decide if a word could be English.

The ERP results might further support this strategy theory, showing a more pronounced N400 component for Hungarian words than for English. The N400 is widely understood as a surprise signal, having higher amplitudes for unexpected stimuli. We propose that the more negative N400 could be a sign of a mismatch between the expected and actual language of an item. Since the homographs could easily be seen as English, they met the expectation criteria, hence the in-between N400 component.

Alternatively, the elevated N400 could also signify richer semantic representations and neighborhoods for Hungarian words. We argue, however, that this is less likely since the homographs had an equally high frequency in the Hungarian corpus as the non-homograph Hungarian words; if the recognition is invariant to language expectation, then these words should also show an N400 at least as prominent as the Hungarian ones.

The lack of any early differences between the ERP waveform shows that the first stages of word recognition do not differ for Hungarian, English, or homograph words, or at least not in this experiment. This might be because both Latin-based scripts require similar processing steps (perhaps N200 differences would arise when comparing alphabetic scripts to syllabaries or left-to-right writing systems to right-to-left ones). The most apparent visual difference between Hungarian and English scripts is the absence of diacritics in the latter. This, apparently, is not enough to elicit a large-scale neural difference detectable with ERP.

Based on the visual word recognition models, the conclusion can be drawn that both lexicons of a bilingual individual are active (Dijkstra et al., 1999). The processing of interlexical homographs confirms that phonological and semantic representations are needed to identify a visual word besides orthographic awareness. In the case of written word recognition, phonological activation occurs, as was previously stated in the semantic, orthographic, and phonological interactive activation model.

For the co-activation of both lexicons, Lemhöfer and Dijkstra (2004) gave the BIA+ model as an explanation. According to BIA+ (Dijkstra & Van Heuven, 2002), the visual presentation of a word leads to parallel activation of orthographic input representations in L1 and L2. These representations activate semantic and phonological representations, resulting in complex code interaction. When the appropriate language gets selected, the input word is recognized. Moreover, BIA+ says that interlexical homographs have separate representations for each language. BIA+ furthermore emphasizes that the activation of various lexical representations is continuously audited by the task/decision system, which supports task execution and decision-making (Green, 1998).

The reaction time of the recognition of homographs is slower for bilinguals since they are exposed to two meanings of homographs. Hsieh et al. (2017) also give the BIA and BIA+ models (Dijkstra & Van Heuven, 1998, 2002; Thomas & Van Heuven, 2005) as an explanation since all nodes between languages are interconnected at the word level, and they mutually inhibit each other. Slower reaction times for interlexical homographs suggest that bilinguals face a competition of representations from their L1 and L2 during the processing of homographs (Hsieh et al., 2017). The data support language non-selectivity, meaning there is an automatic co-activation of information in both linguistic subsystems.

The response time of homographs is also longer because processing printed words continues until the orthographic word unit is recognized and the orthographic representation meets the linguistic properties (phonology, morphology, semantics). According to Carreiras (2013), the boundary line between orthographic and linguistics processing is fuzzy at this point. Nazir et al. (2004) furthermore explain that high-level considerations form the distributional characteristic features of letters in the given language, and the word recognition system learns these properties that make reading successful. Words with high-frequency results in perceptual learning that helps fast and effective word recognition, so word frequency also influences word recognition (Frost, 2012; Kronbichler, 2004). Neurolinguistic evidence (Simos et al., 2002; Solomyak & Marantz, 2010; Szwed et al. (2012) suggests that although high-level linguistic information already exists at approximately 100 ms from stimulus onset, the visual system responds only to the frequency of letter strings, and lexical and phonological features are taken into consideration much later. It also explains why the recognition of cognates and interlexical homographs takes a longer time.

5. Conclusion

The present study provides evidence for co-activation and competition between languages in bilingual word processing. In recognition of homographs, answers indicate a bias towards English responses. The coefficient revealed a high relationship between the ratio of Hungarian replies and the Zipf-frequency difference between Hungarian and English. Multiple comparisons confirmed no difference in the mean response times of Hungarian and English words, whereas homographs produced longer response times. There was no significant difference between the two categories in the early stages of recognition, corresponding with the orthographic-phonological level. This result indicates the relative ease with which the participants can process letter strings from both L1 and L2. However, the brain representations of the two languages diverged later. The ERP waveforms did not demonstrate any significant variations between items regarded as English or Hungarian in the case of the Hungarian-English homographs. Although Hungarian and English have different writing systems and are typologically unrelated languages, the processing patterns are very similar. Although recognizing interlexical homographs does not trigger different processing patterns, various cognitive efforts can be observed according to the decision-making.

Altogether, we could replicate the homograph effect and found that the differences can be at least partly explained by the decision-making strategies of the participants. To test our theories, we propose future experiments to control the strategy by rephrasing the participants' task to concentrate on one or the other language and see if the response bias changes direction.

References

Aitchison, J. (1987). *Words in the Mind: An Introduction to the Mental Lexicon*. Oxford: Basil Blackwell

Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, 53, 60–80 https://doi.org/10.1016/j.jml.2005.02.004

Assadollahi, R., & Pulvermuller, F. (2003). Early influences of word length and frequency: a group study using MEG. *Neuroreport* 14, 1183–1187 https://doi.org/10.1097/00001756-200306110-00016

Beauvillain, C., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language*, 26, 658–672. https://doi.org/10.1016/0749-596X(87)90108-2

Bice, K., & Kroll, J. F. (2015). Native language change during early stages of second language learning. *NeuroReport*, 26, 966–971. https://doi.org/10.1097/WNR.0000000000000453

Braun, M., Jacobs, A. M., Hahne, A., Ricker, B., Hofmann, M., & Hutzler, F. (2006). Modelgenerated lexical activity predicts graded ERP amplitudes in lexical decision. *Brain Res.* 1073–1074, 431–439. https://doi.org/10.1016/j.brainres.2005.12.078.

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2013). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, 18(2), 90–98. https://doi.org/10.1016/j.tics.2013.11.005

Carter, R. (2009). *The Human Brain Book: An Illustrated Guide to its Structure Function, and Disorders.* London: DK Publishing

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first language speech production. *Journal of Phonetics,* 40, 249–268. DOI: 10.1016/j.wocn.2011.10.007

Chang, C. B. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics*, 41, 520–533. https://doi.org/10.1016/j.wocn.2013.09.006

Csépe, V. (2006). Literacy acquisition and dyslexia in Hungarian. In: Joshi, R.M. and Aaron, P.G. (Eds) *Handbook of orthography and literacy*. Mahwah; London: Lawrence Erlbaum Associates. 231-247.

Dambacher, M., Kliegl, R., Hofmann, M., & Jacobs, A. M. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Res*. 1084, 89–103 https://doi.org/10.1016/j.brainres.2006.02.010

De Groot, A. M. B. (2011). *Language and Cognition in Bilinguals and Multilinguals*. New York: Psychology Press.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Bilingualism: Language and Cognition* 5: 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, 41, 496–518. https://doi.org/10.1006/jmla.1999.2654

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language*, 42, 445–464. https://doi.org/10.1006/jmla.1999.2697

Dussias, P. E., & Scaltz, T. R. C. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologica*, 128, 501–513. DOI: 10.1016/j.actpsy.2007.09.004

Elston-Guttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: global language context and adjustment affect processing of interlingual homographs in sentences. *Brain Res. Cogn. Brain Res*. 25, 57–70. https://doi.org/10.1016/j.cogbrainres.2005.04.007

Fabbro, F. (2000) Introduction to language and cerebellum. *Journal of Neurolinguistics*. 13, pp. 83-94. DOI: 10.1016/S0911-6044(00)00005-1

Fabbro, F., & Paradis, M. (1995) Differential impairments in four multilingual patients with subcortical lesions. In M. Paradis (ed.) *Aspects of bilingual aphasia*. Oxford: Pergamon Press. pp. 139-176. DOI: 10.1006/brln.2001.2480

Frost, R. (2012). A universal approach to modeling visual word recognition and reading: not only possible, but also inevitable. *Behav. Brain Sci*. 35, 310–329 doi: 10.1017/s0140525x12000635

Goldrick, M., Runnqvist, E., & Costa, A. (2014). Language switching makes pronunciation less native-like. *Psychological Science*, 25, 1031–1036. https://doi.org/10.1177/0956797613520014

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

Green, D., & Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol*. 25, 515–530. doi: 10.1080/20445911.2013.796377

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1, 131–149.

Grosjean, F. (2001). "The bilingual's language modes," in *One mind, Two languages: Bilingual Language Processing*, ed L. N. Janet (Oxford: Blackwell), 1–22. doi: 10.3389/fpsyg.2010.00178

Heredia, R. R. (1996). Bilingual Memory: A Re-Revised Version of the Hierarchical Model of Bilingual Memory. *The newsletter of the Centre for Research in Language*, University of California, San Diego, 10, January. DOI: 10.1007/978-1-4614-9218-4

Hsieh, M-C., Jeong, H., Kawata, K. H. D. S., Sasaki, Y., Lee, H-C., Yokoyama, S., Sugiura, M., & Kawashima, R. (2017). Neural correlates of bilingual language control during interlingual homograph processing in a logogram writing system. *Brain & Language* 174 (2017) 72-85. http://dx.doi.org/10.1016/j.bandl.2017.06.006

Hull, R., & Vaid, J. (2005). Clearing the cobwebs from the study of the bilingual brain: Converging evidence from laterality and electrophysiological research. In. J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 480–496). New York: Oxford University Press.

Hull, R., & Vaid, J. (2007). Bilingual language lateralization: A meta-analytic tale of two hemispheres. *Neuropsychologia*. 45. pp. 1987-2008. DOI: 10.1016/j.neuropsychologia.2007.03.002

Kerkhofs, R., Dijkstra, T., Chwilla, D. J., & De Bruijn, E. R. A. (2006). Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects. *Brain Research*, 1068, 170–183. https://doi.org/10.1016/j.brainres.2005.10.087

Kim, K. H., Relkin, N. R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*. 388. pp. 171-174. DOI: 10.1038/40623

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence from asymmetric connections between bilingual memory representations. *Journal of Memory and Language*. 33, 149-174. https://doi.org/10.1006/jmla.1994.1008

Kroll, J. F., & De Groot, A. M. B. (2005). *Handbook of Bilingualism: Psycholinguistic Approaches*. New York, Oxford University Press.

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics*, 1, 1–18. doi: 10.1146/annurev-linguist-030514-124937

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. *Neuroimage*. 2004 Mar;21(3):946-53. doi: 10.1016/j.neuroimage.2003.10.021. PMID: 15006661.

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition*, 32, 533–550. https://doi.org/10.3758/BF03195845

Macizo, P., Bajo, T., & Cruz Martín, M. (2010). Inhibitory processes in bilingual language comprehension: evidence from Spanish–English interlexical homographs. *J. Mem. Lang*. 63, 232–244. doi: 10.1016/j.jml.2010.04.002

Malt, B. C., Li, P., Pavlenko, A., Zhu, H., & Ameel, E. (2015). Bidirectional lexical interaction in late immersed Mandarin–English bilinguals. *Journal of Memory and Language*, 82, 86–104. DOI: 10.1016/j.jml.2015.03.001

Marrero, M., Golden, C., & Espe-Pfeifer, P. (2002). Bilingualism, brain injury, and recovery. Implications for understanding the bilingual and for therapy. *Clinical Psychology Review*. 22. pp. 463-478. DOI: 10.1016/s0272-7358(01)00109-x

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, 7, 703–704. DOI: 10.1038/nn1264

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck- Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning*, 60, 123–150. https://doi.org/10.1111/j.1467-9922.2010.00604.x

Mechelli, A., Crinion, J., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R., & Price, C. (2004). Structural plasticity in the bilingual brain. *Nature*. 431. pp. 757. https://doi.org/10.1038/431757a

Navracsics, J., & Sáry, Gy. (2013). Vizuális szófelismerés a kétnyelvűségben. *Alkalmazott Nyelvtudomány* XIII. évf. 1-2. szám 2013.

Navracsics, J. (2007). Word classes and the bilingual mental lexicon. In: Lengyel Zs. és Navracsics J. (szerk.) *Second Language Lexical Processes: Applied linguistic and psycholinguistic perspectives.* Clevedon: Multilingual Matters. pp. 17-39. DOI: 10.21832/9781853599682-004

Nazir, T.A., Ben-Boutayab, N., Decoppet, N., Deutsch, A., & Frost, R. (2004). Reading habits, perceptual learning, and recognition of printed words. *Brain Lang*. 2004 Mar;88(3):294-311. doi: 10.1016/S0093-934X(03)00168-8. PMID: 14967213.

Obler, L. K. (1981). Right hemisphere participation in second language acquisition. In K. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 53–64). Rowley, MA: Newbury.

Pavlenko, A. (2009). Conceptual representation in the bilingual lexicon and second language vocabulary learning. In Pavlenko A (ed.) *The bilingual mental lexicon. Interdisciplinary approaches.* Bristol-Buffalo-Toronto: Multilingual Matters. pp. 125 160.

Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes* 22. 106–129. https://doi.org/10.1080/01690960500463920

Simos, P. G., Breier, J. I., Fletcher, J. M., Foorman, B. R., Castillo, E. M., & Papanicaolaou, A. C. (2002). Brain Mechanisms for Reading Words and Pseudowords: an Integrated Approach. *Cerebral Cortex* Mar 2002; 12:297-305; 1047-3211/02/$4.00. Oxford University Press doi: 10.1093/cercor/12.3.297.

Singleton, D. (1999). *Exploring the Second Language Mental Lexicon*. Cambridge: Cambridge University Press.

Solin, D. (1989). The systematic misrepresentation of bilingual-crossed aphasia data and its consequences. *Brain and Language*. 36. pp. 92-116. https://doi.org/10.1016/0093-934X(89)90054-0

Solomyak, O., & Marantz, A. (2010). Evidence for early morphological decomposition in visual word recognition. *J. Cogn. Neurosci*. 22, 2042–2057 https://doi.org/10.1162/jocn.2009.21296

Studnitz, V. R. E., & Green, D. W. (2002). Interlingual homograph interference in German-English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition*, Roswitha E., 5, 1–23. http://dx.doi.org/10.1017/S1366728902000111.

Szwed, M., Vinckier, F., Cohen, L., & Dehaene, S. (2012). Towards a universal neurobiological architecture for learning to read. *Behav Brain Sci*. 2012 Oct;35(5):308-9. doi: 10.1017/S0140525X12000283. Epub 2012 Aug 29. PMID: 22929623.

Thomas, M. S. C., & van Heuven, W. J. B. (2005). Computational Models of Bilingual Comprehension. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 202–225). Oxford University Press.

Vaid, J., & Genesee, F. (1980). Neuropsychological approaches to bilingualism. *Canadian Journal of Psychology*, 34, 417–445. https://doi.org/10.1037/h0081108

Vaid, J., & Hall, D. G. (1991). Neuropsychological perspectives on bilingualism: Right, left, and center. In A. Reynolds (Ed.), *Bilingualism, multiculturalism, and second language learning*: The McGill conference in honor of Wallace E. Lambert (pp. 81–112). Hillsdale, NJ: Erlbaum. https://doi.org/10.4324/9781315807478

Van Assche, E., Duyck, W., Hartsuiker, R., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science* 20. 923–927.

Van Heuven, W. J. B., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Research Reviews*, 64, 104–122. DOI: 10.1016/j.brainresrev.2010.03.002

Yamins, D. L., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc Natl Acad Sci U S A*. 2014 Jun 10;111(23):8619-24. doi: 10.1073/pnas.1403112111. Epub 2014 May 8. PMID: 24812127; PMCID: PMC4060707.

Yum, Y. N., & Law, S-P. (2021). N170 reflects visual familiarity and automatic sublexical phonological access in L2 written word processing. *Bilingualism: Language and Cognition* 24, 670–680. https://doi.org/10.1017/S1366728920000759

Zatorre, R. (1989). On the representation of multiple languages in the brain: old problems and new directions. *Brain and Language* 36. pp. 127-147. https://doi.org/10.1016/0093-934X(89)90056-4