Answers to Dr. Varga Vera

First of all, I would like to express my sincere gratitude for reading and reviewing my dissertation. I appreciate your kind words regarding the importance and practical relevance of the research topic.

It is true that the Introduction is a bit longer and includes sections about the brain structures that might seem redundant. However, I do believe that there is a need for a comprehensive outline, given that the dissertation's main goal is to give guidance to linguists, language teachers and language learners, who, presumably, do not have the necessary knowledge of the relationship between the brain and language and bilingual visual language processing.

Since I tested the Bilingual Interactive Activation+ model (Dijkstra & Van Heuven, 2002) in my study, I intended to provide an insight into the most relevant and most frequently cited models. I understand that you miss the discussion of the Dual-Route Cascade model (Coltheart et al., 2001). I decided not to include the model since it discusses reading aloud and mostly used in dyslexia research. Also, thank you for highlighting the Bi-Modal Interactive Activation Model (Grainger & Holcomb, 2009), which explains the direction of neighborhood density effects in visual word recognition according to grapheme-phoneme consistency. It would be a great base of further research.

I appreciate that you found the aim of the study, the research questions, and the hypotheses clearly stated and well-motivated.

Thank you for drawing my attention to the difference between phonological awareness and phonological processing, I will use these terms unambiguously in the future. Also, thank you for suggesting that I should test phonological awareness with phoneme deletion tests or rhyme judgment tasks.

I am grateful to you for highlighting your concern about the word superiority effect. I understand that it refers to the phenomenon that letters in words are recognized faster and more accurately than single letters or letters embedded in pseudo-words or non-words. When formulating Hypothesis no. 3 (the recognition of non-words is faster), I meant that for participants, it would be easier to identify nonsense letter strings than to identify words with meaning. I propose BIA+ model as an explanation. In the case of non-words, the recognition can occur at early phases (at as early as 100 ms), since this is the level of orthography and this is where the identification of letter strings takes place. In case of real words, N170 and N400 are also present, which are responsible for phonology, and the lexico-semantic processing, respectively. In this case, the semantics of a word is checked, which takes more time. Although in this lexical decision task, the group mean response times were 649 and 648 ms for the word and non-word conditions respectively, I suggest that carrying out an experiment with pseudo-words and non-words would be beneficial and could result in interesting findings.

I appreciate that you find the methodological details of the experiments thoroughly described, the stimulus lists well-prepared, and the supplements useful to aid understanding the design.

Although I intended to study Hungarian-English bilinguals, I understand that it would have been better to compare the results with a control group, too, and the findings could have been stronger if a monolingual group had been examined. However, it would be quite challenging to find real monolinguals due to the multilingual, dominantly English stimuli around us delivered by the media, social media, movies, music, billboards, etc.

Thank you for drawing my attention to the misinterpretation of lexical decision and language decision tasks. That is right, in the third test, participants were instructed to decide whether the pseudo-words presented on the screen could fit into the Hungarian or into the English language. Still, at the moment I think a test with pseudo-words is rather a lexical decision task than a language decision one. In the future, I will read more about this.

In the first experiment, to compare N400 component amplitudes, I averaged voltage levels in the time window between 380 and 420 ms post stimulus onset. N400 occurs at 400 ms, but it can extend from 380 to 420 ms. I chose channel C1 because N400 is generally maximal over the central (centro-parietal) electrode sites. Furthermore, N400 is associated with lexico-semantic processing, and the central parts of the brain are responsible for semantic processing.

Thank you for highlighting that the results section provides clear and detailed description of the findings, the figures are informative, and the data analysis is impressive.

I appreciate your detailed comments on the issue of shallow and deep writing systems, the interpretation of the N170 component, and the active areas of the brain. I will take every suggestion of yours into consideration in my future career.

Now I attempt to answer your questions.

(1) *What could be the practical application of the results, and what are the limitations of their application?*

Reading requires the perception of the printed word, which is a fundamental ability. Although the identification of printed words is frequently studied in monolingual situations, bilingual written language processing remains an unexplored topic, particularly with Hungarian as a component of bilingualism. At the same time, research on bilingual written word processing can provide crucial information not only for researchers but also for teachers who deal with bi- or multilingual children and facilitate their literacy development. Research on visual word recognition of bilinguals is fundamental, since numerous bilingual students attend monolingual educational institutions, and teachers have to be aware of what is happening in a bilingual student's mind when they are facing reading or writing exercises, since they have to cope with two languages.

Bilinguals keep both their languages active as they use them interchangeably in their everyday lives. The participants in my study have C1 level English proficiency and use English at work and in their everyday lives on a daily basis. The experiments were designed to keep both languages active for this reason. In this way, we can get a realistic picture of how written words are processed. The experiment can indicate the problems of young bilingual readers (as we know the general myth about the "problems" seen by language teachers and speech pathologists in the simultaneous acquisition of reading in two languages).

(2) *Do the results provide support or refute any of the bilingual mental lexicon models?*

The results support the Hierarchical Model of lexical and conceptual representation in bilingual memory (Kroll & Stewart, 1994). The model suggests that fluency must be taken into account, and in the case of less fluent bilinguals, there is a dual store, but in the case of fluent bilinguals, there is a single store conceptual representation, which suggests a common storage of languages. In the present study participants kept both their languages active due to the fact that they were all highly proficient bilinguals. That is why there was high accuracy for both Hungarian and English conditions in the language decision test. Participants did not need any special effort to identify the words, and the homographs were equally processed regardless the language, furthermore, the ERP curves did not represent significant difference either.

The results also provide support for the BIA+ model. For the co-activation of both lexicons BIA+ might be the explanation (Dijkstra & Van Heuven, 2002). The visual presentation of a word leads to parallel activation of orthographic input representations in L1 and L2. Semantic and phonological representations are activated, and it ends up in a complex interaction between codes. When the appropriate language gets selected, the input word is recognized. Moreover, BIA+ says that homographs have separate representations for each language.

(3) *I propose an alternative explanation for the results. The result that only the word vs. non-word contrast yielded a statistically different result in the early time-window (150-300 ms), while the English and Hungarian words and pseudoword differed in the latter time window (mainly in the N400 component) could be explained based purely on orthographic processing without any reference to phonology. The word vs non-word contrast might evoke different N170 response based on their different bigram (especially positional bigram) frequency. Non-word contains letter clusters that are infrequent or nonexistent in a given orthography; thus, purely visual statistical sensitivity to these letter combinations can result in differential modulation of the N170 amplitude. In a similar vein, if the bigram frequencies do not give the reader any clue about what language is being presented (as would be the case if the bigram frequencies were equated between the Hungarian and the English stimuli which might have happened as a "side effect" of equating the word frequencies), no processing difference would emerge until semantic access. Then, if the word has two meanings as a homograph, higher N400 response is recorded than if it has only one meaning. If the words have no meaning but contain existing bigrams (that is they are pseudowords), they are processed similarly to word real words. If the bigram frequencies do not differ between the Hungarian and English pseudowords, there will be no difference in their processing early on that could help them in the decision. However, once semantic processing starts, the reader must realize that the stimulus has no meaning in either language, but have to decide to which language to fit the pseudoword. This decision result in the difference between Hungarian and English pseudowords in the late time window (after 500 ms). This could be based on some decision related revision process that recalculates the bigram frequencies and more precisely compares them in the two languages in order to make the decision about language. What is the Candidate opinion, is any phonological or articulatory processing necessary in this task? How could the two accounts be tested empirically?*

Thank you for the alternative explanation of the results. You are right, only the recognition of words vs. non-words yielded statistical difference in the temporal area in the early (150-200 ms) and late (200-250 ms) phrases of N170 ERP component, which is the perceptual phase of word recognition, which can be explained by the orthographic awareness. Thank you for drawing my attention to the bigram frequencies, it would also lead to interesting findings about neighborhood density, which I did not investigate in my study.

For testing the phonological and articulatory processing, I believe that comparing words and pseudo-words in another experimental design would be a good idea, and I would expect that pseudo-words cause greater activations in certain brain regions than words carrying a meaning (De Groot, 2011; Shaul et al., 2012, Carreiras et al., 2013). This greater brain activity clarifies that unknown stimuli that are incapable of accessing word associations might activate the neuronal network more than words that the individual is already familiar with. Pseudo-words are associated with word-specific mental representations. In the recognition of pseudo-words and words that have rare equivocal orthography-pronunciation correspondence, the lexical representation generates the retrieval of the word. Pseudo-words require a higher-level phonological awareness. On the other hand, for an experienced reader, reading a word carrying a meaning that has a high frequency does not require much phonology, and the recognition does not depend on lexical retrieval, the process is rather automatized. The importance of word frequency was also proved in a study by Simos et al. (2000), in which they gained evidence for activations in different areas depending on the frequency of the word. Perea et al. (2005) examined how frequency influences lexical decision, and confirmed that the frequency of words that are used to create pseudo-words determines how participants recognize them. They found that pseudo-words that were generated by changing one internal letter of the original word, pseudo-words with high frequency showed slower latencies than pseudo-words with low frequency. In the case of high-frequency pseudo-words that were generated by changing two adjacent internal letters, the latencies were also slower than the ones with low frequency. But in the case of one-letter different pseudo-words the high frequent ones showed faster latencies than the ones with low frequency.

Last but not least, I really appreciate your detailed comments and suggestions, they will be helpful in my future career.

Veszprém, 22. 09. 2023

Ihász Petra