Title: SECOND LANGUAGE LEARNERS' STRATEGIES OF READING MULTIMODAL TEXTS: THE EFFECT OF SOCIAL MEDIA USE ON READING

My responses to Prof. Kata Csizér's comments

- **There are some formal inconsistencies and language-related issues, such as occasional short paragraphs, lack of references for statements, inconsistent of italics and absence of indentation.**

Formal inconsistencies and language-related issues will be corrected.

- **It remains unclear why the study is referred to as "pilot research" in multiple sections in the dissertation.**

The current research represents an emerging and relatively uncharted area of study, making it challenging to anticipate the outcomes and potential obstacles. Moreover, the study was carried out during the COVID-19 pandemic, which posed significant difficulties in securing equipment and recruiting participants. Consequently, conducting both pilot and main studies simultaneously, as originally intended, was not feasible.

Thus, my dissertation, is appropriately labeled as a "pilot study" for several reasons. Firstly, it allowed me to gain confidence in the research design. Secondly, it served as a small-scale preliminary investigation aimed at assessing its feasibility and determining the practicality of the research methods. Lastly, as detailed in Chapter 6, this pilot study played a crucial role in refining and optimizing the research design and instruments, as well as identifying potential flaws or weaknesses in the process.

This piece of information will be added to the dissertation.

- **The introduction is organized in a logical way and effectively sets the scene for the research. I miss a more pronounced justification of the apparent research gap and the rationale for the current study, which is clearly adequate for a PhD dissertation. It is somewhat confusing to read Sections 1.2 and 1.3 without references as it is not clear whether the author is presenting her own opinions or referring to previous empirical studies. Many of the ideas are repeated in section 2.10, still without references.**

It's important to note that my own perspectives, derived from years of experience in English language teaching, have shaped the content in section 1.3. and the third paragraph in section 2.10. I will ensure that the text reflects my unique voice and insights. I will adapt the language in these sections to mirror my perspective and insights. In these sections, my primary aim was simply to explain the personal motivation and experience that influenced the selection of this topic for my research. However, I will provide references to the available literature where possible. Some relevant references are provided below:

Amirtharaj, A. D., Raghavan, D., & Arulappan, J. (2023). Preferences for printed books versus E-books among university students in a Middle Eastern country. *Heliyon*, *9*(6), e16776. https://doi.org/10.1016/j.heliyon.2023.e16776

Bennett, S., Maton, K., Kervin, L. (2008). The 'digital natives' debate: A critical review of the evidence. *British Journal of Educational Technology*, *39*(5), 775-786. https://doi.org/10.1111/j.1467-8535.2007.00793.x

Carr, N. G. (2011). *The shallows: what the Internet is doing to our brains*. Norton pbk. [ed.] New York, W.W. Norton.

Gündüz, U. (2017). The effect of social media on identity construction. *Mediterranean Journal of Social Sciences*. *8*(5), 85-92.

Hayles, K. H. (2007). Hyper and deep attention: The generational divide in cognitive modes. *Media*, *1*, 187–199.

Manalu, H. B. (2019). Student's perception of Digital texts reading: a case study at the English education department of Universitas Kristen Indonesi. *Journal of English Teaching*, *5*(3), 191-203. https://doi.org/10.33541/jet.v5i3.1312

Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International journal of educational research*, *58*, 61-68. https://doi.org/10.1016/j.ijer.2012.12.002

Owusu-Acheaw, M (2016). Social media usage and its impact on reading habits: a study of Koforidua Polytechnic students. *International Journal of Social Media and Interactive Learning Environments (IJSMILE)*, *4*(3), 211-222. https://doi.org/10.1504/IJSMILE.2016.079493

Ramamohanarao, K., Gupta, K. K., Peng, T., & Leckie, C. (2007). The Curse of Ease of Access to the Internet. In P. McDaniel & S. K. Gupta (Eds.). *Information Systems Security*. Lecture Notes in Computer Science (Vol. 4812). Springer. https://doi.org/10.1007/978-3-540-77086-2_18

- **The literature review summarizes the most important theories and empirical studies pertaining to the research topic, overviewing multimodal reading and age-related differences. The structure is logical, and the subchapters and sections are effectively signposted. My overall comment here concerns the fact that the review is mainly a summary of previous studies without achieving a good balance between description and evaluation. More specifically, I miss the critical overview of previous theoretical and empirical studies, which would help justifying the research gap in a valid way.**

Thank you for your valuable feedback on the literature review section. I appreciate your acknowledgment of the logical structure and effective signposting within the subchapters and sections. I understand your point regarding the need for a more critical evaluation of the previous theoretical and empirical studies. I will take this into consideration to enhance the review by incorporating a deeper analysis that goes beyond mere summarization. I will revise the literature review to strike a more balanced approach between description and evaluation, ensuring a clearer justification for the research gap. Your input has been incredibly helpful in guiding the refinement of this section.

- **The methods sections lack several pieces information, including details about the population/populations of the study as well as some aspects of the sampling methods.**

I believe the aspects of sampling methods are discussed in detail in sections 3.4 and 3.5. Section 3.4 explores how the age of second-language learners affects their reading habits and multimodal reading. It involved two groups: 20 secondary school students selected through a B1 Euroexam practice test, and 61 language school students chosen based on a placement test and subsequent interviews assessing their speaking skills. For the secondary school students, the Euroexam was administered online due to COVID-19 restrictions.

Only 9 out of 20 students met the B1 proficiency level in reading. The language school students underwent a different selection process involving a written placement test and speaking evaluations. In this section, I've also outlined the reasons for choosing the B1 English language proficiency level and the Euroexam test.

Section 3.5 explains the eye-tracking experiment which utilized technology to study how L2 learners read multimodal texts. All the participants completed the B1 Euroexam practice test, achieving pass marks. Due to COVID-19 restrictions, non-random sampling methods (convenience and voluntary response) were employed. Calls for participation were made on university Facebook pages, resulting in 15 volunteers. Ten achieved pass marks, but only seven had usable data due to technical issues with recordings/calibrations for three participants. Additionally, two students from the language school I worked for voluntarily participated, completed the Euroexam test via email and obtained pass marks.

Regarding the population details, I acknowledge the oversight in not providing comprehensive information in my initial submission. I will include a detailed description of the studied populations to offer a comprehensive view in future discussions.

- **It is unclear why some levels of measurement were simplified in the study (e.g., age) and why a categorical variable was approximated with a normal distribution.**

The concept of measuring age is explored in detail in sections 3.2.1, 3.6, and 4.2.2. These measurement levels were originally designed to discern significant differences among age groups, enhancing our understanding of the subject. Initially, the investigation did not include multimodal reading in different age groups as a research focus. Due to challenges in recruiting participants within specific age categories, I collected data from my students, who spanned various age ranges. Consequently, I revised the research questions, adjusted the subject parameters, and enlisted students from the language school where I worked, a group I could easily access. Given my familiarity with my students, I categorized them into two groups: those likely to be part of an educational environment and those who were less likely. It is important to acknowledge that for more robust future research, information on the students' educational status should be gathered to confirm their academic setting.

However, because this research question emerged and evolved during the study, it was not initially considered at the outset.

In our study, I simplified certain levels of measurement, including age, to enhance the feasibility of data analysis and interpretation. I approximated a categorical variable with a normal distribution as a simplifying assumption for statistical analysis. This decision was made after careful consideration of the research objectives and the available data. First, for the purposes of our analysis, it was more practical to group individuals into specific age categories to simplify the interpretation and to make the results more accessible to a broader audience. Besides, grouping ages into categories can enhance statistical power, especially if the sample size is limited. Analyzing age as a continuous variable might require a larger sample to detect meaningful differences. Simplifying the levels of measurement can also aid in the interpretation of the results, making them more understandable and applicable for practical decision-making. Moreover, categorical variables are approximated with a normal distribution to allow the use of parametric statistical tests (e.g., ANOVA) for analysis when the data distribution is not severely skewed. This approach has also been chosen to better align with the research questions and practical applications of the study.

However, I did perform non-parametric tests (e.g., Mann-Whitney U, Kruskal-Wallis or Chi-Square goodness of fit) as an alternative analysis, and these tests yielded results consistent with our parametric approach, further supporting the validity of our findings.

I acknowledge the trade-offs involved in simplification and the potential loss of information. I will include a discussion in the revised manuscript highlighting these limitations and considering potential future research directions that might involve more detailed age analysis.

- **Additionally, it is also unclear how sample sizes were determined.**

At the beginning of the eye-tracking experiment, the optimum number of participants was set at 30, as suggested by Pernice and Nielsen (2009). However, the COVID-19 pandemic presented considerable challenges in terms of participant accessibility, leading to the inability to achieve the intended sample size. Nevertheless, even with this reduced sample

size, valuable insights into multimodal reading were gained, given that the primary focus of the eye-tracking experiment was on meaning rather than formulating generalized hypotheses (Crouch & Mckenzie, 2006).

Similarly, for the online experiment, the initial plan was to gather a minimum of 30 participants (Yıldırım & Şimşek, 2006; Baykul, 1999; Ross, 2004).

The determination of sample sizes also adhered to the guidelines provided by Cochran (1977) and Krejcie and Morgan (1970). Based on these guidelines, a population of 92 individuals participating in the online experiment should have justified a sample size of 75. However, due to the pandemic-related constraints, only 70 participants were ultimately collected.

- **Concerning the online questionnaire, there is a lack of information about quality control measures implemented to ensure that the instrument collects reliable and valid data. While there is information about the design of the test, there is a missing link to explicit references justifying both the design choices and its overall quality. Essentially, we know what the candidate did, but we lack insight into why those specific actions were taken.**

To ensure the questionnaire's solid theoretical foundation, we incorporated Martinec and Salway's (2005) as well as Engebretsen's (2012) frameworks, both extensively expounded upon in chapters 2 and 3. These frameworks serve as the underpinnings of our research, guaranteeing that the questionnaire's questions align with theoretical principles and are pertinent to the study's objectives.

Additionally, while the questionnaire was rooted in theoretical constructs, I selected its elements in a way that allowed a decent evaluation of the research question's vital aspect within the study's limitations. This involved a consideration of key variables and constructs (e.g., visual-verbal relations; visual and verbal preferences/responses) essential for addressing our research aims.

Furthermore, we conducted statistical analyses to provide insights into the instrument's reliability and validity. As detailed in chapter six, we assessed internal consistency using Cronbach's alpha coefficient, which yielded a value of 0.70, signifying the test items'

reasonable reliability and their contribution to overall consistency. To evaluate test validity, a Pearson Correlation analysis was carried out to determine the relationships between each item in the multimodal reading test and its overall score. These results substantiate the test's accurate measurement of the intended construct, confirming that the questions effectively capture the research's targeted aspects.

As emphasized in chapter 6, although the multimodal reading test demonstrated satisfactory reliability and validity, these aspects alone do not guarantee practical significance. Consequently, I took a critical stance to identify the shortcomings of my questionnaire and I delved into the challenges encountered during the research in chapter 6, exploring avenues for enhancing the test's applicability. These practical challenges indicate my commitment to further developing the test to overcome these issues.

- **The follow-up questions appear to lack complexity, I am not sure about the quality of the data collected with them.**

While the questions may appear straightforward and lack intricacy, they have provided significant insights into the modes that enhanced the participants' comprehension and their ability to answer the questions. In the future research, more intricate questions could be included as follow-up inquiries to obtain a more comprehensive understanding of readers' modal preferences. For example, after the participants have read and answered the questions for each multimodal text, the follow-up questions could be repeated to help explore how the combination of image and text influences understanding of each specific multimodal text.

Some other possible questions could be:

Did the pictures:

a) Make the text clearer.

b) Add more information.

c) Confuse the meaning.

d) Were not helpful.

Which was more important for understanding:

a) The text.

b) The pictures.

c) Both equally.

d) Neither helped.

When the text and the picture were different:

a) I followed the text.

b) I followed the picture.

c) I combined both.

d) I ignored both.

When pictures were unclear:

a) I ignored them.

b) I searched for more clues.

c) I relied on the text.

d) I assumed they weren't important.

- **I think the candidate uses the words population and sample interchangeably, which complicates understanding some of her points.**

Your comment is right. To address this concern, I will make a clear differentiation between the population and the sample. The population is referred to as the entire group that is the subject of study or about which conclusions are to be made. While a sample is a subset of the population that is studied to gather information and draw inferences about the larger population. I will provide explicit definitions and the information about the population to clarify these distinctions and ensure consistency in their usage throughout the dissertation.

- **The use of the word "experiment" requires clarification, as essential components like treatments and pre- and post-test measures or control group are absent.**

In my research, the term 'experiment' is used to describe an ongoing and evolving investigative process rather than a controlled experiment with fixed components such as treatments, pre- and post-test measures, or control groups. My work involves a dynamic approach that incorporates various adjustments and adaptations as part of its methodology. The research design is centered on a continuous and flexible exploration of the subject

matter, which allows for the incorporation of new elements and insights as the study progresses. I will add this explanation to my work to ensure the clarity of the term.

- **Unfortunately, tables do not contain sample sizes, making it challenging to assess the potential levels of sampling error.**

Your comment is right. I will add sample sizes to the tables.

- **Test-values are often missing from the tables, and, in some instances, the measurement units are unclear. The level of significance cannot be zero, I am not sure what the zeros mean in the text and in some of the tables.**

The reported P-value of .00 should not be misinterpreted as zero. The apparent value of zero comes from the limitation of representing results with only two decimal places in my research, but it does indeed exceed zero.

I will use a standard notation, e.g. $p < .05$ or .01, .001, .0001, .00001 depending on the p value.

I'll also include Test-values and the measurement units in the table to clarify the information.

- **Furthermore, it is unclear what motivated the use of both parametric and non-parametric statistical procedures.**

In this dissertation parametric statistical procedures are mainly used with a couple of exceptions. One of the primary reasons for using both parametric and non-parametric tests is that our dataset contains a variety of variables with different data distributions. We aimed to accommodate the diversity of our data and ensure that the statistical methods were appropriate for each variable.

Moreover, I chose to employ a combination of parametric and nonparametric tests because certain variables in our study did not meet the assumptions required for parametric tests, necessitating the use of nonparametric alternatives. Additionally, certain variables, such as nominal variables, inherently call for nonparametric tests due to their categorical nature.

- **I genuinely appreciate the candidate's effort to discuss her results meaningfully and in a comprehensive manner. It demonstrates that she invested time in engaging with her findings and connecting them back to the reviewed**

literature. There seems to be a good balance between summarization and discussion, although I do miss the candidate's voice. It would be interesting to know which results surprised her most and why. Personally, I found Chapter 6 to be surprising after the discussion, I think information presented there is more suitable for the methods section.

I did not include my voice because my intention was to maintain objectivity and avoid introducing a strong personal voice due to the nature of the topic, which allows for various interpretations. Nonetheless, if I were to provide a more personal perspective, the analysis of data from the eye-tracking experiment was quite illuminating. It underscored how seemingly minor factors such as color, text complexity and visual density, among others, might significantly influence the comprehension of multimodal texts. My research has only scratched the surface of a rapidly growing field of research, and there are numerous avenues for further exploration and investigation.

This arrangement provides insights into the limitations of my research instruments subsequent to the extensive process of data analysis. By positioning Chapter 6 in this specific location, I aim to serve a dual purpose. Firstly, it enhances the understanding of potential constraints associated with the research instruments, which became evident through an extensive review of the data. Secondly, this placement is a reminder to the readers, emphasizing the need for a cautious interpretation of the results in light of the identified limitations that have surfaced over the course of the study.

However, given your comment, I will relocate the discussion on validity and reliability to the methods section.

The conclusion of the dissertation covers all the necessary elements, including limitations, future research directions, and pedagogical implications. While I would have appreciated a more synthetic approach to future research directions, it is commendable that they are linked to the apparent shortcomings of the study.

On the basis of the above considerations, I can conclude that this thesis presents an original piece of research on a timely and important topic, and, therefore, I think that

it meets the requirements of a PhD degree. I am looking forward to discussing the dissertation with the candidate at the public defence.

Thank you for your invaluable contribution as a reviewer of my PhD dissertation. Your meticulous examination, insightful comments, and constructive feedback have been immensely beneficial in enhancing the quality and depth of my research.

References:

Baykul, Y. (1999). *İstatistik metotlar ve uygulamalar (3. bs)*. Ankara: Anı Yayıncılık.

Cochran, W. G. (1977). Sampling techniques (3rd ed.). New York: Wiley.

Crouch, M., & McKenzie, Heather (2006). The logic of small samples in interview based qualitativeresearch. *Social Science Information, 45*(4), 483-499.

Engebretsen, A. (2012). Balancing cohesion and tension in multimodal rhetoric. An interdisciplinary approach to the study of semiotic complexity. *Learning, Media and Technology, 37*(2), 145-162. DOI: 10.1080/17439884.2012.655745

Krejcie, R. V. & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement, 30*, 607-610.

Martinec, R., & Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication, 4*(3), 337-3.

Pernice, K., & Nielsen, J. (2009). *How to Conduct and Evaluate Usability Studies using Eyetracking*. Nielsen Norman Group: Fremont, CA, USA.

Ross, M. R. (2004). *Introduction to probability and statistics for engineers and scientists.* San Diego: Elsevier Academic Press.

Yildirim, A., & Simsek, H. (2008). *Nitel arastirma yontemleri [Qualitative research methods]*. Seckin.