A review of Sheida Marzban's Ph.D. dissertation *Second language learners' strategies of reading multimodal texts: the effect of social media use on reading***

My responses to Dr. Márta Lesznyák's comments

General remarks

The aim of the dissertation is to "investigate reading habits, modal preferences, and multimodal strategies among different age groups of second language learners (L2).". The topic is timely as multimodal text processing has become an everyday reality. In the context of foreign language learning, course books and online materials contain lots of multimodal elements that learners have to process. Visual information is believed to support learning, but it also leads to changes in the reading process as it puts an extra load on cognitive processing. The dissertation reports the results of a questionnaire- and an eye-tracking study. The research design is sound, but because of time constraints the author could only complete the pilot study, as a result, findings of the research can hardly be generalized. The structure of the dissertation follows what is expected of empirical research studies. The author uses robust statistics to analyze her data. In my opinion, however, it is not the results that constitutes the most valuable part of the dissertation, but the effort to create a novel questionnaire on multimodal reading. Although the questionnaire that was used for data collection, suffers from numerous shortcomings, the author reflects on these and offers suggestions for improvement. This is very promising: on the one hand, it shows that the candidate is creative and able to come up with novel research ideas, to design her own instrument and reflect on it and improve it. On the other hand, it also suggests that the questionnaire can really be turned into a valid and reliable instrument that can be used in future research, too.

In detail:

- **After an Introduction in Chapter 1., the theoretical review is presented in Chapter 2. First, the concept of multimodality is defined and described, then models which explain interaction between various semiotic modes (i.e. visual and textual) are discussed. In Section 2.3 possible relations of text and picture are reviewed, paying special attention to cohesion and tension. Section 2.4 focuses on multimodality and learning in general, whereas section 2.5 is devoted to multimodality in second language learning, specifically. In this section, some course book analyses are reviewed, and this is followed by the overview of research on using caption (subtitles) in videos for language learning purposes. Finally, research on reading tasks with static visual information is discussed. In these sections, a large amount of information is provided, nevertheless, no attempt is made to highlight what is relevant for this research and how.**

I believe the sequence of the subchapters creates a cohesive journey through various aspects of multimodality which provides the necessary theoretical scaffolding and sets the stage for its application within the context of this research. However, I acknowledge your comment regarding the need for greater focus and relevance to the current study. I will include explicit connections between the reviewed literature and its relevance to my research in future discussions to demonstrate the applicability of the literature to the study.

- **The conclusions of Section 2.9.1 (p. 42) sound somewhat unfounded, as the chapter is about studies on multimodal processing in social media and nothing is written about school policies. Then the author concludes that schools are more open to phones. This may be true but seem to be unrelated to the previously presented literature.**

Your comment is right and I will remove this information from the conclusion.

- **Section 2.9.2 focuses on the effect of social media on Academic Reading habits. However, throughout the chapter, academic reading habits are not defined, so it is not clear what the author is writing about. When presenting previous research – again, it is not described how academic reading habits were operationalized. Reading habits are defined in the Methodology section on page 52, but as this is a key term, it should have been defined earlier and more clearly.**

In my research, academic reading habits are defined as the patterns, behaviors, and approaches individuals adopt when engaging with texts and materials for educational purposes. This includes but is not limited to the frequency and consistency of engaging with educational materials. Operationalizing academic reading habits in research also involves translating the abstract concept of "academic reading habits" into measurable and observable variables that can be studied, analyzed, and quantified. These could include factors such as the frequency and duration of reading materials, the types of materials read and reading strategies employed, to name a few. In order to asses these factors, measurement tools and instruments should be developed. This might involve designing surveys, questionnaires, or structured interviews specifically tailored to gather data on reading behaviors, preferences, and strategies used by individuals in educational settings. This information will be added to the dissertation accordingly.

- **The author does not provide a general overview of reading, thus the review is slightly mosaic-like: some important aspects are missing. Thus, the author does not deal with the topic of text types, and the fact that different text types may include different amounts of visual information, as a result, different strategies might be needed to process them. There seems to be an overtone that suggests that multimodal reading is the norm nowadays. This might be true for everyday personal communication, but it is hardly the case in any sort of professional communication, where sometimes you meet 100% verbal texts, other times multimodal processing has long been the norm (i.e. figures and diagrams, X-rays, drawings etc.). It is not emphasized that the**

multimodal processing studied here is just one part of the overall reading skills (even multimodal reading skills) needed in the 21st century.

I understand the concern regarding the absence of a comprehensive overview of reading, particularly in relation to text types and the varying amounts of visual information they may contain. However, I decided not to include a literature review on different text types and associated reading strategies in my dissertation because my research focus didn't center on multimodal reading across various text types. Each text type incorporates distinct levels of visual information, leading to differing reading approaches. Yet, delving into reading strategies for diverse text types (e.g., advertisements, webpages, PowerPoint presentations) goes beyond the current scope of my research goals. However, exploring multimodal reading strategies for different text types would certainly be a promising avenue for future research.

Moreover, I didn't mean to imply that multimodal reading is the norm nowadays. I acknowledge that while multimodal reading is on the rise in personal communication, it might not hold the same prominence in certain professional settings where purely verbal texts or specific visual aids, such as figures and diagrams, are quite common. My primary focus on multimodal reading only stems from the dissertation's requirement for an in-depth review of this particular topic.

- **The author does not take into account the context and the aim of the reading either. Multimodal processing can be highly dependent on what the aim of the reading task is. The author unconsciously taps into this problem by overviewing reading in learning (languages). However, learning is not the only aim of reading, there can be several others: e.g. entertainment, information extraction, information exchange, special professional (reading for translation, proofreading, reading for evaluation, reading for summarizing etc.). From this perspective, it would have been a good idea to contextualize the reading tasks on the test.**

Regarding your insightful comment on contextualization and reading task objectives, one of the focal points of this dissertation involves exploring the reading habits of L2 learners to assess the potential impact of social media on them. Social media platforms often engage users in aimless scrolling across various pages and channels. Consequently, no specific instructions were provided to the participants regarding the purpose of the multimodal reading test. Their task solely involved reading the multimodal texts and responding to related questions. This approach was adopted to try to simulate the freeform reading experience similar to social media. The participants were encouraged to navigate fluidly between text and image and answer the questions. However, investigating the influence of different reading objectives on multimodal reading processes presents an intriguing avenue for future research. Such exploration could significantly contribute to the advancement of this field of study.

In light of your feedback, I will incorporate this information into my work to enhance the contextualization of the reading tasks in the test.

- **Reading habit questionnaire (Page 52.) – how was the questionnaire developed? Did you rely on previous questionnaires and literature? If so, what were these?**

I designed the questions by synthesizing information from various scholarly sources related to reading habits (See below for references) and also incorporated my own insights into the mix. Given the scope of the study and its preliminary nature, I attempted to make a balance between depth and conciseness in the questionnaire design. The intention was to create a concise yet effective tool that could provide insightful data without overwhelming the participants, especially considering it wasn't the primary instrument of the study and it was done exactly prior to the main trial. To develop the questionnaire, a review of existing literature on reading habits and related questionnaires was conducted. Drawing from both the insights from the literature review and my interactions with students during class discussions, a collection of possible questionnaire items was formulated. These items were designed to capture the frequency of reading different forms of media. Subsequently, I narrowed my focus to inquiries around primary media prevalent in our daily lives, including books, magazines and social media platforms. Given my primary emphasis on multimodality and social media, additional statements on various modes (text, image and video) within social media were also incorporated. The questionnaire initially in English was translated into Hungarian by a native Hungarian speaker.
I will add this information to the future discussions.

Mirza, Q., Pathan, H., Khatoon, S., Hassan, A., (2021). Digital age and reading habits: empirical evidence from Pakistani engineering university. *TESOL International Journal*,*16* (1), 210-136.
Owusu-Acheaw, M. (2016). Social media usage and its impact on reading habits: a study of Koforidua Polytechnic students. *Int. J. Soc. Media Interact. Learn. Environ.*, *4*, 211-222.
Qenaway, Kh. (2019). The impacts of using social media on second language reading habit among Arab students. Academia. Retrieved from https://www.academia.edu/40216569/the_impacts_of_using_social_media_on_second_language_reading_habit_among_arab_students.
Rafiq, M., Khan, T. M., Asim, A., & Arif, M. (2019). The effects of social media on reading habits. *Pakistan Journal of Information Management & Libraries, 21*, 46–65.

- **The multimodal reading test is a completely novel test developed by the candidate. The idea of the test is great, and it is on its way to become a good one. Nevertheless, I would think that the way the test is structured, there is no room for real multimodal processing, or at least, the respondents have no chance to show that they engaged in multimodal processing. This is because they are forced to choose between 2 options, there is not "both" or "neither" option. Using open-ended questions would have offered the opportunity to participants to indicate spontaneously how they processed the text and the visual information.**

I acknowledge the point you raised regarding the restricted options within the test format of the online experiment, limiting participants' ability to express engagement in multimodal processing adequately. To address this limitation, I incorporated an eye-tracking experiment, aiming to delve deeper into multimodal processing regardless of the participants' responses to the questions.

However, your suggestion about incorporating open-ended questions and more options (neither; both) is insightful, providing respondents with the opportunity to spontaneously articulate their approach towards processing textual and visual information. Utilizing open-ended questions could also enhance the test's effectiveness by allowing the participants to express their thought processes more freely. This modification not only enables a deeper understanding of how individuals engage with multimodal stimuli but also aligns with the dynamic nature of multimodal processing. I acknowledge that your input is incredibly instrumental in improving the current test structure and will pave the way for future refinement and development of the multimodal reading test.

- Moreover, the test items right now do not seem to reliably assess modal preference and only modal preference. Some items lend themselves to several interpretations, others are contradictory or confusing. Nevertheless, the author reflects on these problems in the Discussion section and offers solutions for them, which is really promising but does not help the interpretation of the results received with this version of the questionnaire.
When the multimodal reading test is discussed, it is not indicated whether the author alone decided about the types of image-text relations and other features of the items. If so, it would be beneficial to involve other experts in making such decisions.

I agree that there might be confusion with some test items. Although the Discussion section offers solutions, I recognize that the current questionnaire might not show these fixes, affecting result interpretation. Understanding this drives me to refine the questionnaire in the future for a clearer assessment of preferences.

I made decisions on the test's design, but your idea to involve other experts is important. Working with diverse experts could improve the test's accuracy by bringing in different viewpoints. Including a group of experts from relevant areas to refine the test should be considered in the future research. This could enhance the questionnaire's reliability and validity in future research to make sure it accurately measures preferences without confusion.

- **Concerning data collection procedures I have two questions: why were the participants not allowed to use the “back” button? And why they were not informed about this?**

The rationale behind disabling the "back" button during the study was aligned with one of the primary research objectives: investigating participants' modal preferences. The multimodal questionnaire incorporated visual and verbal responses at the end of each multimodal text, to capture the participants' immediate response. Additionally, the follow-up questionnaire, provided at the end of the multimodal reading test, aimed to investigate which mode (image, text, or both) facilitated their reading and response processes. Consequently, disabling the "back" button prevented the participants from revisiting the multimodal text to enhance recall, ensuring that their initial, spontaneous responses were captured, which was a focal aspect of the investigation.

The absence of explicit information provided to the participants regarding the disabled "back" button was intentional. I intended to avoid influencing participants' behaviors towards memorization of the multimodal text. My goal was to elicit immediate, unfiltered responses, emphasizing the authenticity of their initial reactions rather than encouraging deliberate recall.

- **On page 63, Table 6 shows the differences between the online reading test and the eye-tracking test, but no justification or explanation is given for the changes.**

The adjustments made in the test aimed to optimize its efficacy while aligning with the research objectives. Specifically, 'Craigdarroch Inn' was substituted with 'George Inn' due to readability concerns and the resulting extended reading duration, which did not align with the intended research focus. The replacement with 'George Inn,' an easier and shorter term, was important in ensuring a smoother reading process.
Similarly, the alteration from 'He is putting something on the desk' to 'He is putting it on the desk' was made to enhance the authenticity and reduce potential confusion within the reading context. This adjustment aimed to maintain coherence and clarity within the text, ensuring that the task reflected a more natural and understandable sentence structure for participants.

- **For the future: it would be interesting to do the same test with a control group in their 1st language (Hungarian in this case).**

I genuinely appreciate your feedback and the insightful suggestion. It is indeed compelling and presents an intriguing avenue for further exploration. The inclusion of a control group using their primary language could offer valuable insights into potential differences in multimodal processing between languages.

- **Methods of data analysis are described clearly, although no information is given on how data on reading- and response duration was collected. (With what software?)**

In order to collect data in the online and eye-tracking experiments, Flexiquiz platform and OGAMA software, as mentioned in chapter 3, were used respectively. These tools allowed me to gather information regarding the time duration spent on each slide, including both text reading and response slides.

Results:

The results are presented in an appropriate style and manner, and the statistical tests employed are adequate. Nevertheless, sometimes the formulation of the results is slightly misleading:

- **e.g. p. 72-73. "*The results show 65.60% of the participants acknowledged that the pictures helped their reading comprehension ('Yes' condition) while 34% of the participants did not find the pictures helpful ('No' condition). According to the Chi-square test results, the relationship between the two variables was also significant (X2***

***(1, N = 61) = 61.00, p= .00).*" – Do you mean that there is a significant difference between the "yes" and the "no" answers?**

Your comment is right. The Chi-square test shows a significant difference between the 'Yes' and 'No' responses concerning how participants perceived the usefulness of the pictures in enhancing their reading comprehension. This information will be corrected.

- **Page 77-78.: apparently, the author uses noteworthy as a synonym for significant, which is inaccurate if we talk about statistics. Significant difference means that the difference is real (= it is not a measurement error). Non-significant differences should not be interpreted as difference, as they can be measurement errors. (This paragraph suggests that the author is not aware of this).**

You've rightly pointed out that in statistical contexts, "significant" indicates a real difference beyond measurement error, while "non-significant" differences should not be interpreted as difference but possibly as a result of measurement errors. I acknowledge the importance of accurate terminology in statistical analysis and I will correct this piece of information (significant instead of noteworthy) accordingly.

Discussion:

In the Discussion section the author provides insightful interpretations of her findings. Below, I will make some comments on some of her explanations.

- **p. 96. "*This could be due to the fact that what schools offer to students and prepare them for is quite different from the visual and pictorial world outside school* (Kress, 2003)." – The term „outside school" is vague: it includes domains that are visual and domains that are verbal, and domains that are both visual and verbal. The school as an academic domain can't do otherwise than rely on verbal processing. But it does not mean that students are not able to rank visual information first or to integrate it with verbal information if it is needed. This is something we do not have information about.**

Your comment raises an essential point about the term "outside school". I aimed to convey that the education system often emphasizes verbal processing, which might not fully align with the visual and pictorial elements prevalent in our everyday life beyond educational settings. You rightly note that the term "outside school" includes various domains, including visual, verbal, and those combining both. Indeed, while the academic setting predominantly relies on verbal processing, it doesn't inherently imply that students are incapable of prioritizing visual information or integrating it with verbal content when necessary. I'll ensure to consider these dimensions in my future work.

- **p. 96. „*As a consequence, it may have hindered the participants from adopting a multimodal approach when reading multimodal texts*." – It is not clear what the author means by multimodal approach and how she can decide whether the subjects have adapted it or not? The *process* itself was not studied. The question arises whether processing both visual and verbal information, *then deciding to rely on one of them* is a multimodal strategy or not. Also, when answering the follow-up questions, the respondents expressed a positive attitude towards the visual information presented and found them useful. This indicates that they did process this information, whatever answer they gave. My question to you would be: What result (answer) would indicate for you that the subject has engaged in multimodal processing?**

Thank you for your valuable insight. The sentence you're referencing comes after this statement, "The presence of questions in the multimodal reading test may have inadvertently influenced participants' responses, possibly creating a washback effect similar to that observed in academic assessments conducted in schools and universities (Alderson & Wall, 1993)". I aimed to express that the presence of the tests/questions and the requirement to respond to them could have impacted their performance, encouraging a preference for verbal over visual responses. The way the sentence is currently written doesn't accurately convey this meaning and should be rephrased to better express this intent. While I believe multimodal processing might refer to the processing and utilization of both visual and verbal information while engaging with multimodal texts, I would like to clarify that investigating and assessing multimodal processing wasn't the central aim of the online experiment. Therefore, this information will be corrected accordingly.

- **p. 106. Calculating Cronbach alpha and Pearson r for the questionnaire pre-supposes an item-based scoring of the multimodal reading test. However, the scoring system is not described in the dissertation, although it can be figured out how it must have been done.**

The scoring system employed for the multimodal reading test involved assigning scores to each individual item based on the participants' responses. Each visual and verbal response was assigned a score of 1. The total number of visual and verbal responses selected by the participants were counted at the end of the test. I will ensure to include a clear and explicit description of the scoring system applied to the multimodal reading test in my dissertation

- **A more serious problem is presented by Table 23 (page 107) which claims to show the Pearson correlation coefficient "between each question in the multimodal reading" test. Looking at the format of the table, this can hardly be the case. I would think that the table shows each item's correlation with the total test score.**

The title of Table 23, "The results of the Pearson Correlation between each question in the multimodal reading test and its overall score," intended to show the correlation between each individual item within the multimodal reading test and the test's overall score. To ensure clarity, the intention behind the table's content was repeated in Section 6.1, paragraphs 2 and 3, emphasizing the correlation between individual test items and the total score.

- **Moreover, while item-total correlations can be a good indicator of construct validity, there are other types of validity that are not dealt with at all. The critical one would be content validity, that is, whether the test really assesses multimodal processing (and not some other cognitive processes).**

Thank you for highlighting the significance of content validity to assess whether the test genuinely measures multimodal processing or other cognitive processes. I agree that while item-total correlations provide insights into construct validity, content validity holds substantial importance to evaluate multimodal processing.
Although my dissertation may not have explicitly delved into content validity to examine whether the test measures precisely the targeted construct of multimodal processing, this vital aspect should be addressed in future iterations of the research.

- **In the subsections of Chapter 6.2, critical items are reviewed and modifications in the wordings, positionings of texts and visuals, choice of visuals are suggested. These suggestions show that the candidate is able to analyze her instruments and findings critically. However, we do not learn how the author realized that these items are controversial, confusing etc. Was it on the basis of test scores? Did another expert look at the items? Did she herself notice these problems?**

Thank you for your insightful comments regarding section 6.2. You rightly pointed out that while I suggested modifications to critical items in my study, I didn't explicitly elaborate on how I identified these items as controversial or confusing. I gained a deeper understanding of these issues through the process of submitting articles based on the findings of my dissertation to various journals and conferences. The reviews I received on these articles shed light on potential shortcomings and areas needing improvement in my questionnaire. This external feedback from experts and peers played a significant role in highlighting concerns regarding the wordings, positioning of texts and visuals, and the choice of visuals within my questionnaire. Upon reflecting on the feedback received from these conferences and journal reviews, I developed a more critical perspective towards my questionnaire. This critical evaluation led me to re-examine and reassess specific items within my instruments, identifying those that might be contentious or ambiguous. I'll ensure to include this process of gaining insight from external reviews and conferences in future discussions and provide a clearer understanding of how I recognized certain issues within my questionnaire.

Conclusions

In this chapter, the candidate summarizes the main findings of the research, discusses the limitations and the implications of the study and gives suggestions for future research. “The Limitations of the study” shows that the author understands the shortcomings of her research and will be able to avoid these problems in the future.

- **In the "Implications of the research" sections the author offers some suggestions for textbook developers and teachers. These suggestions are relevant but it is not clear how they come from the results of this research.**

The suggestions put forth for textbook developers and educators were provided with the intention of bridging the gap between the research outcomes and their practical applications. These implications primarily stem from the insights gained through the eye-tracking experiment conducted during this study (e.g., different intermodal interactions and degrees of redundancy in textbooks, duplications and semantic gaps in apps and webs). While the implications were informed by the findings and insights emerging from the research outcomes, especially the data obtained through the eye-tracking experiment, I acknowledge the need for a more explicit and transparent link between specific results and the suggested implications for textbook developers and teachers in future discussion.

- **Moreover, it can be somewhat risky to propose that teaching should be adjusted to students' preferences: whereas it can be fun for teenagers to read texts from social media in some classes, they must get acquainted with other genres and text types, too.**

Thank you for your insightful comment. I completely agree that while integrating texts from social media or other preferred sources might engage students and add an element of enjoyment, a well-rounded education necessitates exposure to diverse genres and text types.

The proposal to adapt teaching methods to students' preferences was not intended to advocate exclusivity the use of certain types of texts but rather to underscore the importance of acknowledging and incorporating students' interests into the learning process. It's crucial to find a balance between using preferred materials, like social media texts, to keep students engaged and exposing them to various genres and text types for their overall development.

In future discussions, I will emphasize the importance of this balanced approach, highlighting the significance of integrating diverse texts while incorporating students' preferences to create an engaging learning environment.

- **Making textbooks and assessment more multimodal in schools is another difficult issue. Whereas it is relatively easy to visualize cognitively easy material for beginners or young children, it becomes increasingly difficult as the curriculum content is getting more abstract. This is something you discussed in the theoretical review, too.**

Indeed, the process of integrating multimodal content becomes notably more complex as the curriculum progresses from simpler, more concrete concepts to abstract, higher-level content.

As I also mentioned this as one of the difficulties in my questionnaire development in chapter 3, the visualization of straightforward material for beginners or younger learners tends to be relatively easier. However, translating more abstract or complex curriculum content into multimodal formats poses considerable challenges. This complexity emphasizes the necessity for innovative approaches and careful considerations in designing multimodal elements that accurately represent and enhance understanding, particularly within advanced curriculum content.

- **Also, English textbooks (particularly language coursebooks) are usually full of visuals. Very often, a larger proportion of a page is taken up by photos and pictures than by texts.**

Indeed, English textbooks often allocate a substantial portion of pages to images, creating a visually rich learning environment. I would like to clarify that in the implications section, my intent was not to negate the presence of images in textbooks but to underscore the significance of material and textbook developers, web and app designers, and teachers in critically analyzing the existing multimodal materials used in educational settings. It's crucial for these stakeholders to assess whether these materials effectively cater to the current needs and preferences of students.
Moreover, I emphasized the importance of investigating the optimal combination of semiotic modes—visual, textual, and others—to enhance student engagement, interpretation, and attention. While acknowledging the prevalence of visuals in textbooks, the emphasis was on the strategic use of multimodal elements to foster deeper engagement and comprehension among students. By encouraging a critical assessment of available multimodal resources, my aim was to highlight the need for educational stakeholders to be more innovative in their approach.

- **As for assessment, visuals are often avoided because a person must be able to interpret a text without visual assistance. It is very often the case in real life too, you cannot avoid that.**

I completely agree that assessments usually emphasize text interpretation without visual assistance, mirroring real-world situations where such skills are crucial. However, advocating for multimodal assessments doesn't imply that every assessment should exclusively rely on images. Rather, it suggests the value of introducing multimodal assessments that resemble the diverse modes of communication, including visuals, encountered in language learners' daily lives, especially in the realm of social media. Balancing text-based assessments with occasional multimodal assessments presents an opportunity for students to demonstrate comprehension and communication skills across various mediums. This approach not only accommodates different learning styles but also prepares students for the multimodal communication landscape in today's world. Integrating multimodal assessments might complement text-based assessments which is likely to foster a more comprehensive evaluation and cater to the diverse learning styles of students.

Formal aspects:

The dissertation fulfils the formal requirements. Nevertheless, there are minor language issues, like unusual word choices and collocations, subject-verb agreement mistakes, capital letters in the middle of a sentence etc.

Your comment is right. Language issues will be corrected.

Summary

The topic of the dissertation is highly relevant and the theoretical overview of the field is excellent. The candidate developed a multimodal reading questionnaire, with the help of which she could collect data and more importantly, she could reflect on the properties of the questionnaire and

suggest modifications for it. Coupled with eye-tracking, the questionnaire helped the candidate uncover cognitive processes of subjects who were involved in multi-modal reading processes.

In consequence, I recommend that the dissertation is presented at the oral defence . If the oral defence is successful, I recommend the award of a Phd-degree to the candidate.

I extend my heartfelt thanks for your invaluable review of my work. Your insights and feedback have been immensely helpful. I truly appreciate your time and expertise.