# MICHIGAN STATE
## UNIVERSITY

Karolina Baranowska Dissertation, External Examiner Report
22 July 2022

This is an outstanding dissertation that investigates a topic of current relevance and interest from an interesting theoretical and methodological perspective. The main topic of the effects of subtitling on second language (L2) learner comprehension and acquisition of vocabulary while watching English videos and movies is of particular importance as the access to English expands worldwide. The topic is especially important for countries such as Poland where, a decade or two ago, access to English media was much more limited. In fact, the dissertation has implications for other languages as well because now learners of almost any L2 can find videos and movies in the target language. Therefore, it is important for (a) second language acquisition (SLA) researchers to understand the theoretical issues associated with this learning opportunity , and (b) L2 teachers to understand how L2 movies and videos might best be incorporated into their curriculum.

Before beginning a more detailed assessment of the dissertation, I would like to say that it was a pleasure to read. The dissertation is very well-written and well-organized. The structure is appropriate for a quasi-experimental study, beginning with a review of the literature, followed by an explanation of the methods and then presentation of results. Finally, the dissertation ends with a discussion of the results. The manuscript was excellently edited and proofread, and I was able to find only two or three typos throughout, which is unusual for a document of this length.

**College of Arts and Letters**

Second Language Studies

Michigan State University
619 Red Cedar Rd.
B331 Wells Hall
East Lansing, MI 48824

517-432-1812
Fax: 517-432-2736
http://sls.msu.edu

As for the sections themselves, the literature review consists of three chapters, each reviewing an important focus of the dissertation. Chapter 1 reviews the use of foreign language movies and subtitles in English language teaching. In this chapter, the candidate provides a comprehensive review of the role of L2 movies, especially in the European context, exploring the different policies and practices that various countries have adopted pertaining to subtitling and dubbing movies. The chapter identifies the different options that are available for subtitling, such as using the first language (L1) or L2, and then reviews studies that have researched this topic. The chapter also presents the pedagogic importance of the movies and subtitling. As an instructed SLA researcher myself, I always appreciate it when research studies have direct implications for what teachers do in their classrooms.

Below are a couple of comments based on specific parts of Chapter 1.

p. 15-16, The dissertation states that 'watching FL videos compensates for the limited contact with native speakers and their culture'. I think this is an excellent

point, and I would expand on it a bit by saying that FL videos are especially useful for the development of L2 pragmatics (i.e., the use of language in social contexts), especially for students in English as a foreign language (EFL) context in which the only social context for L2 interaction is the classroom which provides a very limited range of social contexts.

p. 28, There is mention of a 2008 study that states that 'subtitles detract from the ecological validity of video transmissions…'. I would like to see a bit more critique of that position from a 2022 perspective in which subtitles are available in numerous real-world contexts. The ubiquity of the availability of subtitles makes this dissertation all the more relevant, and I don't think it hurts to emphasize the changes that have occurred in the last twenty years regarding the availability of technology for L2 learning. In addition, the use of subtitles increases in L1 contexts as educators, and others, are concerned with issues of equity and accessibility. Best practice now recommends that subtitles be included on all videos so that deaf and hard of hearing individuals can access the video material.

Chapter 2 provides a cognitive theoretical support for the dissertation in the form of Cognitive Load Theory, which is an excellent lens through which to view the issue of subtitling for L2 learning. Cognitive Load Theory is described very clearly, and its application to the current study is made obvious. This chapter, in particular, is written in very technical language, but it is still presented in an understandable and comprehensible manner. In addition to presenting the dissertation's supporting theoretical perspective, this chapter also makes a strong argument for the methodological techniques that are used in the study.

Chapter 3 addresses the issue of modality, which is another primary concern of the dissertation. In particular, Dual Coding Theory, is explored to consider the influence of video and audio, separately and combined. In addition, this chapter explores taxonomies of comprehension which again provide support for methodological investigation of the theoretical construct.

Within Chapter 3, there is a section on incidental and intentional vocabulary learning, which is a pertinent issue when it comes to watching L2 movies. I would encourage the candidate to explore this topic a bit more, especially in relation to the construct of incidental learning. Technically, incidental learning has been used to mean learning without intention; however, in many studies to date, intention hasn't been measured, and instead, studies have relied on Hulstijn's definition of incidental learning occurring when learners are unaware of a subsequent vocabulary test. However, this condition is better referred to as incidental learning conditions because it is still possible for learners to intend to learn vocabulary, even if they are just watching a movie for entertainment.

On page 104 there is mention of 'short-term and long-term memory'. Short term memory is not a construct that's been introduced in the dissertation; rather, the distinction has been made between working memory and long term memory. I'm fairly certain that short term memory and working memory are not being equated here. In fact, I think what's being referred to is short term learning and long term learning, which are indeed terms that occur later. Thus, the use of the term here should either be clarified or removed, with the latter being my recommendation.

Throughout the three lit review chapters, there is extensive discussion of previous studies, which is good and appropriate. In addition, there are several tables that summarize the relevant details of these studies. There are also paragraph length descriptions of most of these studies, and one of my suggestions would be to synthesize these studies more, rather than summarizing them separately. Certainly some of the more important ones might warrant an entire paragraph, but others might not.

Chapter 4 presents the research questions, methodology and the results. Again, these sections are presented clearly and appropriately, with the results containing numerous graphs that help illustrate the results. The research questions are generally clear and concise, with the main questions being divided into sub-questions and with each sub-question being accompanied by a hypothesis.

p. 119, A discussion of and justification for the sample size is necessary. The nine treatment groups have ~ 20 participants each. Are these groups large enough to conduct robust inferential statistics? In addition to justification in the method section, there should be mention of the small sample size in the limitations section.

p. 127, Were the assumptions of ANOVA and regression investigated and met? This information should be reported.

Descriptive statistics (mean, standard deviation, 95% confidence intervals) should accompany each of the inferential statistical analyses. I would prefer to see the descriptives in the results section (4.6 and following); however, I realize that the candidate might not wish to clutter up the text with so much additional information, especially since the information is somewhat recoverable from the Figures that are already in place. Thus, an appendix with all of the descriptive statistics would be an acceptable alternative.

The line graphs are an excellent way to illustrate the ANOVA results. I would, however, recommend having the Y-axis include all possible scores for each instrument that is being analyzed, which means that most of the Y-axes would start at zero, and they would end with the maximum possible score. This way the reader can better place participants' mean scores within the larger context of the instrument in question. For example, in Figure 2, it seems that participants in the

No subtitle/Video condition did very poorly on the comprehension test because their data point is at the bottom of the graph. However, when looking more closely at the Y-axis, it appears that even the No subtitle/Video condition scored over 50% on the comprehension test. This performance would be evident if the entire Y-axis went from 0 to 28.

p. 129, I was confused by the use of the term 'main effects analysis' in the sentence below Figure 2 because what is listed are interaction effects.

p. 133, The correlation matrix of the various levels of the comprehension test did not reveal many strong correlations. What does that say about the validity of the taxonomy that was used? Wouldn't we expect to stronger relationships among the levels, especially if they are hierarchical? Some critique of this issue would be appropriate.

p. 140, I'm curious about the vocabulary loss analysis. On page 126, it states, 'vocabulary loss was calculated by subtracting the results of the immediate post-test on vocabulary from the delayed one to see how many words were lost on their way to long-term memory'. By this calculation, it would seem that better participant scores on the immediate posttest (which is a good thing) could result in negative results when it comes to examining vocabulary loss. For example, if one group averages 6 on the immediate posttest and averages only 2 on the delayed posttest, then they have a vocabulary loss score of 4. If another group scores 0 on the immediate posttest (which is bad) and scores 0 again on the delayed posttest, they have a vocabulary loss of 0, which would seem good. It would be helpful if the candidate clarified this issue.

Chapter 5, Discussion
This chapter does a good job of relating the research study's findings to the various theoretical issues that were outlined in the lit review chapters. As I write below, I think there could be a bit more general critique of Dual Coding Theory based on the study, but the candidate does point out the specific instances in which the results are at odds with the theory.

I like the conclusion on p. 145, 'subtitling…delivered some language, which is always better than no language at all'. Throughout the dissertation, I wondered about the video only condition. It seems like viewers might understand some of the plot by just watching the video without any audio, but it seems unlikely that they would learn any vocabulary. I'm wondering if the candidate included this condition primarily for symmetry or if they feel like this might be a viable option in the L2 classroom. This issue is also addressed on p. 152 when it's stated that the Polish subtitle only group would not be expected to show any vocabulary development because they had no input in English, the target language.

Throughout the chapter, there is discussion of how the results often did not support the Dual Coding Theory. Does the candidate have any critique of the theory to offer? Might there be some more general statements to make in addition to pointing out the specific instances in which the research results didn't support the theory?

On p. 146, I'm wondering if there might have been a ceiling effect such that there were no significant differences between the FULL and AUD condition. In other words, the tests might have been too easy to capture any differences between the two conditions. Of course, it's possible that 'the audio-only format is not inferior', but it might also be worth considering if the tests didn't give the participants the opportunity to differentiate themselves based on the treatment conditions.

The second paragraph of the Conclusion chapter does a nice, succinct job of summarizing the main points of the study:
  - Viewing English videos had a positive effect on Polish L1     speakers
  - English subtitles were found to assist with vocabulary learning
  - Polish subtitles assisted with comprehension
  - Subtitles did not overload working memory capacity
  - Viewers can effectively process information presented via different
     modalities

In summary, I found this an interesting, well-written, and well-conducted dissertation. In my evaluation, I have included specific points for consideration (primarily when the candidate goes on the submit manuscripts for publication), Nevertheless, I have an overall positive assessment of the dissertation, and I think the candidate can proceed to the final stages towards earning a PhD title.

Sincerely,
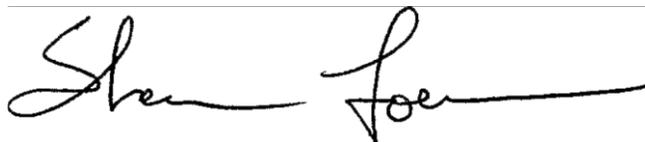
Shawn Loewen, PhD
Professor
Second Language Studies
MATESOL Program
loewens@msu.edu