



Subtitling, Working Memory, and L2 Learning: A Correlational Study

Legendagem, memória de trabalho e aprendizagem em L2: um estudo correlacional

Rafael Matielo*

*Universidade Federal do Ceará (UFC), Fortaleza, Ceará / Brasil
rafaelmatielo@yahoo.com.br

Roberta Pires de Oliveira**

**Universidade Federal de Santa Catarina (UFSC), Florianópolis, Santa Catarina / Brasil
rapiolive@gmail.com

Luciane Baretta***

***Universidade Estadual do Centro-Oeste (UNICENTRO), Guarapuava, Paraná / Brasil
barettaluciane@gmail.com

ABSTRACT: This study investigates the impact of intralingual and interlingual subtitles on Brazilian English as a Foreign Language (EFL) learners as a result of their processing of a North-American sitcom. More specifically, it examines whether subtitling interacts with one's individual differences, working memory (WM) as the case in point. Thirty-six intermediate-level EFL learners were evenly divided into two experimental groups (intralingual subtitles and interlingual subtitles) and one control group (no subtitles). Participants' performance was measured based on an L2 video comprehension test and an L2 vocabulary test. Participants' performance was correlated with their scores on two WM tests. The results obtained revealed that both participants' L2 video comprehension, as well as their L2 vocabulary test performance, did not significantly interact with their WM capacity under any of the experimental conditions. These results are discussed in light of the possible processing mechanisms employed by the participants that may account for the lack of statistically significant correlations found.

KEYWORDS: subtitled videos; L2 video comprehension; L2 vocabulary learning; working memory.

RESUMO: Este estudo investiga o impacto de legendas intralinguais e interlinguais em aprendizes brasileiros de Inglês como Língua Estrangeira (ILE) resultante do processamento de um *sitcom* Norte-Americano. Mais

especificamente, ele examina se a legendagem interage com as diferenças individuais do aprendiz, neste caso, a memória de trabalho (MT). Trinta e seis aprendizes de ILE foram igualmente divididos em dois grupos experimentais (legendas intralingual e legendas interlinguais) e um grupo controle (sem legendas). O desempenho dos participantes foi medido por um teste de compreensão do vídeo em L2 e um de vocabulário em L2. Ademais, o desempenho dos participantes foi correlacionado com os resultados de dois testes de MT. Os resultados obtidos revelaram que a compreensão do vídeo em L2 pelos participantes e seu desempenho no teste de vocabulário em L2 não interagiram significativamente com sua capacidade de MT em nenhuma das condições experimentais. Esses resultados são discutidos à luz de diferentes possíveis mecanismos de processamento empregados pelos participantes que possam explicar a falta de correlações estatisticamente significativas encontradas.

PALAVRAS-CHAVE: vídeos legendados; compreensão de vídeo em L2; aprendizagem de vocabulário em L2; memória de trabalho.

1 Introductory Remarks

With the emergence of Karen Price's seminal work in 1983, a substantial body of knowledge of the effects of and the effects with subtitles in language learning has been gathered, and throughout the last thirty years or so, researchers have moved towards a deeper understanding of how learners may benefit from subtitling in terms of their language development. Despite what may seem to be a long period of academic inquiry, much is still underexplored (VANDERPLANK, 2015).

Investigating the use of intralingual (same language) and interlingual subtitles (different linguistic pair in the audio/subtitles) in L2 learning has been gaining prominence in recent years. From the 1980s on, more than sixty papers have been published in respected journals around the world, reporting on the results of experiments with different populations and target languages (MATIELO; D'ELY; BARETTA, 2015). Interestingly, only a handful of them have addressed Brazilian learners of English (MATIELO; COLLET; D'ELY, 2013; MATIELO; OLIVEIRA; BARETTA, 2017; MATIELO; OLIVEIRA; BARETTA, Forthcoming).

In short, subtitles have been found to foster L2 development, regardless of whether they are interlingual or intralingual (D'YDEWALLE; VAN DE POEL, 1999; HUANG; ESKEY, 1999; KOOLSTRA; BEENTJES, 1999; MARKHAM, 1999; MARKHAM; PETER, 2003; DANAN, 2004; STEWART; PERTUSA, 2004; TAYLOR, 2005; CAIMI, 2006; CHANG,

2006; VAN LOMMEL; LAENEN; D'YDEWALLE, 2006; SYDORENKO, 2010; WINKE; GASS; SYDORENKO, 2010; ZAREI; RASHVAND, 2011; RAINE, 2013). Nonetheless, some studies have not found significant differences considering subtitles' availability and their relationship with the specific language component being tested, such as general/listening comprehension (BIANCHI; CIABATTONI, 2008; LATIFI; MOBALEGH; MOHAMMADI, 2011; MONTERO-PEREZ; PETERS; DESMET, 2013; SHARIF; EBRAHIMIAN, 2013; MATIELO; COLLET; D'ELY, 2013; MATIELO; OLIVEIRA; BARETTA, 2017) and L2 vocabulary learning (MATIELO; COLLET; D'ELY, 2013; ZAREI; GILANIAN, 2013; MATIELO; OLIVEIRA; BARETTA, Forthcoming).

When adding the simultaneous processing of audio and subtitles to the equation, the picture becomes blurrier. Since reading subtitles is considered an automatic process, irrespective of one's familiarity with them or knowledge of the target-language available in the soundtrack – evidence comes from eye-movement recordings (D'YDEWALLE; GIELEN, 1992; WINKE; GASS; SYDORENKO, 2013) – part of one's attentional resources when watching a subtitled film is allocated in the tasks of reading and processing the subtitles, watching and attending to the story, processing the motion picture, and arguably attending to the auditory input. Hence, the ability to successfully attend to the whole set of input – the simultaneous use of spoken (audio/soundtrack) and written (interlingual and intralingual subtitles) input modes – could be linked to one's Working Memory (WM) capacity, an integrated system in charge of temporary storage and manipulation of information during the execution of any given cognitive task (BADDELEY, 1992; BADDELEY, 2011).

Since attending to the whole set of input in a video watching task can be considered highly cognitively demanding, we posit that the systematic exploration of the correlation between learners' WM capacity and their ability to understand the content of the video and pick up novel words encountered in subtitled video materials seems to be relevant and necessary. Insights into this matter may provide L2 researchers and practitioners with further understanding of the challenges that subtitled video materials may pose for L2 learners' simultaneous processing of all the channels at play while engaged in watching subtitled video materials in and outside their L2 classroom environment.

This article is divided into four sections in addition to this introductory one. Section 2 presents the main findings related to subtitling availability and L2 learning from a cognitive oriented perspective. Section 3 centers on the methodological aspects informing this study. Section 4 focuses on the descriptive statistics, the results, and the discussion of the main findings. Finally, section 5 offers a summary of the main findings of the present study while outlining some of their implications.

2 Subtitling and Cognitive Aspects in L2 Learning

Until the 2000s, intralingual subtitles were mostly found to be more effective in aiding L2 development (see also Mاتيело, D'Ely & Baretta, 2015 for a state-of-the-art review), since groups were compared to and contrasted with controls (with no subtitles available in the experiments). As for comparative studies that emerged from the 2000s on, the situation is not quite clear and presents grayer areas. For instance, in two studies (STEWART; PERTUSA, 2004; HAYATI; MOHMEDI, 2011), better results were mostly achieved with intralingual subtitles, whereas in one of the studies (MARKHAM; PETER, 2003), better results were achieved with interlingual subtitles. Some studies presented better results with one or the other depending on the language component being tested or proficiency group (MARKHAM; PETER; MCCARTHY, 2001; BIANCHI; CIABATTONI, 2008; LATIFI; MOBALEGH; MOHAMMADI, 2011; ZAREI; RASHVAND, 2011). As far as video comprehension and vocabulary learning are concerned, there does not seem to be an agreement as to which type of translational aid can be more beneficial.

The findings originated from eye-tracking studies on the processing of subtitled materials point to a few poignant aspects. Perego, Del Missier, Porta, and Mosconi (2010) examined the hypothesis that the processing of subtitled film is cognitively effective, that is, one should be able to understand the film content without a significant tradeoff between the processing of images (visual input) and text (written input). When watching subtitled material, “attention needs to be flexibly allocated on parallel information sources during this task” (p. 250). They hypothesized that when attention is more focused on the subtitles, image processing would be less effective; hence, the opposite should also be true. They explain that this hypothesis is generally consistent with attentional theories that postulate the early

selection of information channels. Results from the study, however, revealed no tradeoff effects between subtitle processing and image processing, thus suggesting that the participants watching the subtitled film did process its content and subtitles effectively.

Kruger and Steyn (2014) offer novel insights from an experiment carried out to investigate subtitle reading behavior and performance. In short, the results obtained with eye-tracking experiments showed that although no significant differences were found between the performance of the participants in the intralingual subtitles condition and those in the control condition, participants in the test group who actually did read the subtitles performed better on the comprehension test than those who saw the videos with the subtitles but did not read the subtitles.

Winke, Gass, and Sydorenko (2013) also looked at caption-reading behavior based on eye-tracking methodologies. Their study's results revealed that, in general, participants fixated on the intralingual subtitles 68% of the time when the captions were shown on the screen. There was also an effect of video (video familiarity) on caption viewing depending on the L2 being learned. More specifically, learners of Arabic, Russian, and Spanish spent similar amount of time reading the captions on both videos. The Chinese learners, in contrast, spent less time reading the captions when watching the video with familiar content.

Based on the findings of these studies, it is clear that learners watching subtitled video materials distribute their attentional resources into processing both the visual scenes and the subtitles. Bird and Williams (2002) explored a central prevailing issue regarding the speculations of the effects of subtitling on L2 learning, that is, whether or not soundtrack is in fact processed when subtitled videos are watched. Their study examined the effects of single modality input – either sound or text – and bimodal input – sound and text – presentation on word learning, with an explicit focus on L2 word learning. Measures consisted of enhancements in spoken word recognition efficiency (implicit memory) and recognition memory related to word retention (explicit memory). Results indicated that auditory lexical decisions in terms of familiar words were equally primed by prior bimodal and sound-only presentation modes, though no priming effects for nonwords were found. Furthermore, the results suggested that the participants were able to attend to and process both text and sound. Nonetheless, bimodal input failed to show any significant advantage in relation to sound-only modality. This is

an extremely important finding that lends support to the notion that the addition of text was not conducive to better gains in learning in relation to single-modality input.

In light of these findings, one's individual differences – WM, more specifically – may play a decisive role in one's successful distribution of attentional resources while processing the different sources of information at the same time: soundtrack, motion picture, and subtitles. It could also be expected that the higher one's WM capacity is, the better one should do on subtitled video materials' processing and language learning tasks. This is particularly pertinent if we take into account that WM directly relates to the control of specific cognitive mechanisms, such as attention, processing, and other regulatory functions, entailing the access to long-term memory information (BADDELEY, 2000).

According to Baddeley (1992), WM research has mostly developed under two different, but complementary approaches. The first approach is the dual-task and neuropsychological approach, focusing on the analysis of the structure of WM itself, emphasizing its slave subsystems, including the study of evidence of neuropsychology and the application of dual tasks, requiring participants to memorize and store digits in a digit-span task while performing other cognitive tasks, for instance. The second approach, called the psychometric correlational, refers to the correlation between individual differences in WM and the performance of cognitive abilities (DANEMAN; CARPENTER, 1980). The methodology of this approach involves tasks and correlates performance in these tasks with performance in other high cognitive tasks. The present study is carried out within the psychometric approach to the study of WM and its possible relationship with other cognitive tasks, namely language learning tasks.

3 Method

This mixed design (qualitative and quantitative) study¹ examines whether one's WM capacity interacts with the processing of intralingual and interlingual subtitled material by Brazilian English as a Foreign Language (EFL) learners.

¹ For the purposes of this article, only results of correlations among WM tests, the L2 video comprehension test, and the L2 vocabulary test are reported.

3.1 Participants

A total of 36 Brazilian intermediate EFL learners, ranging from 18-60 years of age, participated in the data collection (20 female and 16 male) and were chosen based on their proficiency level. Participants were Brazilian Portuguese native speakers enrolled in level 5 (intermediate) in the Extracurricular² (non-credit) Language Courses at Universidade Federal de Santa Catarina (UFSC), Florianópolis, SC, Brazil. Studies on the effects of subtitled video materials have typically involved intermediate learners of the language, and the assumption behind that is that these participants are usually at a threshold proficiency level that enables them to read the subtitles on screen in the foreign language, given their short display time (2-4 seconds, only).

Participants were randomly assigned to one of the treatment groups or the control group: Intralingual Subtitles Group (n = 12), Interlingual Subtitles Group (n = 12); and Control Group – no subtitles (n = 12). Participants received explanations about all the stages in the data collection in the first meeting, when they were invited to participate in the research; signed the Informed Consent Form³; and had the opportunity to clarify doubts related to the goals and the design of the study. However, participants were not told about specific details that could bias their answers during any of the stages in the data collection. The participants' teachers were also given a questionnaire in Portuguese in order to provide more information about their education and experience with English language teaching.

² The Extracurricular (non-credit) Language Courses at UFSC are open to undergraduate and graduate students enrolled at UFSC or any other higher education institution in the area, as well as to faculty members and members of the community. Students can enroll twice a year and can take a placement test to determine their proficiency level in any of the language courses offered: English, French, German, Italian, Japanese, Portuguese as a Second Language, and Spanish. In relation to English courses, language instructors consist of undergraduate and graduate students taking the Letras program, who are supervised by two coordinators.

³ The research project was submitted to the university's Ethics Committee and an approval was obtained, which is logged under code number 36597314.9.0000.0118. Participants' teachers also signed a consent form.

3.2 The TV Series

The participants watched a 20-minute long episode of the American sitcom *The Big Bang Theory*, which premiered in 2007. In Brazil, the show is broadcast with Portuguese subtitles on Warner channel and is a critically acclaimed show. The sitcom (situational comedy) depicts Leonard Hofstadter and Sheldon Cooper, two brilliant physicists who are best friends and roommates. They are also friends with their co-workers Howard Wolowitz, a mechanical engineer, and Rajesh Koothrappali, an astrophysicist. The gang spends their time working on their individual University projects, playing games, watching comic-con related movies or reading comic books. When Penny, the girl next door moves in the neighboring apartment, the story takes a turn as Leonard is intent on trying to go out with her.⁴

The sitcom was selected based on a series of criteria. First of all, it has been used in previous studies (MATIELO; COLLET; D'ELY, 2013; MATIELO; OLIVEIRA; BARETTA, 2017). Additionally, a sitcom was thought to be appealing and appropriate for the target audience. From the profile questionnaire administered in the very first session in the data collection, 31 out of the 36 participants reported watching sitcoms, thus suggesting their familiarity with the genre. The episode used in the study was “The Grasshopper Experiment”, the eighth episode in the first season. The episode was chosen, as it contained a complete storyline and did not require students to be familiar with the series or previous episodes.

3.3 The Video Comprehension Test

Two parts for this test were designed. For the general comprehension part, three ‘why’ questions were posed and participants were instructed orally to answer them in Portuguese or in English and were assured that grammatical errors would be disregarded. To successfully answer the general comprehension questions, participants needed to understand the story and how it unfolded throughout the episode. All of the events in the episode were highly intertwined, which meant that participants had to have some level of general comprehension as regards the connections among the goings-on in the episode to be able to answer the questions. An example of such a question is: “At the end of the episode, why is Sheldon singing and

⁴ Information retrieved from <http://www.imdb.com/title/tt0898266>.

playing?” In the specific comprehension part, participants had to judge five statements about the story narrated on screen on whether they were true or false. The statements essentially provided details about the characters, their relationships, likes and dislikes, attitudes, and information about the story. An example of such statement to be judged as true or false is: “() Raj is used to drinking”.

3.4 The L2 Vocabulary Test

The L2 vocabulary test contained three parts: a pre-test, a test, and a post-test. The pre-test assessed participants’ previous knowledge of the target vocabulary, in the first session. They received a 20-word list in English in which 10 of them were distractors and were asked to write their meaning, a synonym or an explanation in Portuguese or English using their own words. The target words were chosen taking into account factors influencing word learnability (LAUFER, 1997); for instance, some words were chosen because of their facilitated word learning aspect, such as familiar morphemes (e.g., *pointless* and *membership*), whereas others were chosen for their difficulty-inducing factors, such as the presence of foreign morpheme (e.g., *obnoxious*), and some were selected as neutral, such as those related to concreteness or abstractness (e.g. *wrath*).⁵ Moreover, the words chosen are related to the themes portrayed in the episode, but they are not semantically related (ERTEN; TEKIN, 2008).

Regarding the number of times the target-words appear in the selected episode, half of them were uttered and shown in the subtitles – both intralingual and interlingual – twice (*slot*, *membership*, *guinea pigs*, *showdown*, and *wrath*), whereas half was uttered and shown in the subtitles once (*embodiment*, *pointless*, *pusby*, *obnoxious*, and *resemblance*). The short exposure to the input is acknowledged, though the video length is short (20 minutes).

Other important criteria considered in this study relate to whether the words actually appeared in the interlingual and intralingual subtitles and their relevance to the story, which could facilitate participants’ processing and future recognition. Distractors, on the other hand, contained likely familiar vocabulary based on semantic familiarity (e.g., *affection* and *mint*) and unfamiliarity (e.g., *award* and *moisturizer*), taking into account their proficiency level. Word frequency was not controlled.

⁵ All of the examples have been taken from the actual test.

The L2 vocabulary test required participants to analyze 10 target-vocabulary words from the pre-test in English (excluding the distractors) and write their meaning, a synonym, or an explanation in Portuguese or English, using their own words immediately after watching the video, in the second session. In the third session, for the post-test, participants were provided with a test identical to the test to check if they were able to recognize the words they had encountered by the time the test was administered. The participants were asked to write their meaning, a synonym, or an explanation in Portuguese or English using their own words. As in the L2 vocabulary test, distractors were not included.

3.5 The WM Tests

The Reading Span Test (RST) was originally designed by Daneman and Carpenter (1980) to explore the relationship between individual differences in WM capacity of native speakers of English and their comprehension of sentences in addition to the recall of the last words of a group of sentences. In the present study, the Brazilian version of the adopted RST was a modification of the one used in Torres (2003), which was administered in Portuguese (participants' L1) to avoid confounds with participants' L2 proficiency level and to avoid floor effects (scoring too low) due to task difficulty. This study's version of the test comprised 42 unrelated sentences, ranging from 13 to 17 words in length, presented in sets of 2, 3, 4, and 5 sentences.

Two important modifications in the test were made as a result of informal piloting. This study's version of the test did not use sets containing 6 sentences because all of the five participants in the informal piloting could not remember the words when sets of six sentences were shown on screen, neither in the correct order nor out of order of presentation. Additionally, at the end of the test, these participants from the informal piloting reported finding the test too long and too tiresome. Another important modification was the color of the computer screen: instead of using a white color, this study's version of the test used a black color on the screen, since four out of the five participants in the informal piloting reported being extremely uncomfortable with a white screen during the entire practice and testing sessions, and suggested a black screen instead. The font color was adapted to the screen color as well, therefore using white instead of black to suit the

background of the slides (see also MATIELO, 2016, for more information on the test administration and scoring).

As for the Operation Word Span Test (OSPAN), the test was originally designed by Turner and Engle (1989) to investigate the hypothesis that WM capacity is not language specific and can thus be generalized to any cognitively complex task, since it has been suggested to be a reliable measure of WM capacity (CONWAY *et al.*, 2005). The test essentially consists of asking the participant to solve simple mathematical operations while trying to recall a set of unrelated words.

The Brazilian Portuguese version of the test used in this study was designed by Prebianca (2009) based on Turner and Engle's (1989) test. The words in the test were disyllabic, unlikely to be unknown by native speakers of Portuguese in the age range under study (e.g. *papel* and *tinta*). The OSPAN was also administered in Portuguese to avoid confounds with participants' L2 proficiency level. As in the RST, the OSPAN consisted of 42 operation strings along with Portuguese words, written in white and placed directly in the middle of a black computer screen. From the 42 trials, 19 strings presented correct operations, whereas 23 strings displayed incorrect operations (see also MATIELO, 2016, for more information on the test administration and scoring).

Both the RST and the OSPAN tests were scored strictly and leniently. For the strict scoring of RST test, participants' reading span was calculated at the level at which s/he was accurate on at least two trials of a given set of sentences (DANEMAN; CARPENTER, 1980). In accordance with Turner and Engle (1989), an approximately 85% accuracy rate was required in terms of participants' judgment of sentence grammaticality or syntax acceptability to ensure the processing component of the task, which represented 36 out of 42 sentences. Half a point was given when the participant passed one trial at a certain level. For instance, when a participant correctly recalled all the words in the right order in the three sets of 2 sentences and correctly recalled just one trial in the three sets of 3 sentences, this participant received half a point, her/his span being considered 2.5, which is where scoring would then terminate. In the lenient scoring of the RST, participants were given points for any set for which s/he recalled all final sentence words, regardless of the order of recall, as long as s/he recalled all of the words belonging to two of the trials in the given set, while a half point was given when the participant passed one trial at a certain level, which is where scoring

would then terminate. The approximately 85% accuracy rate was also kept in the lenient scoring of the RST test in order to maintain the processing component of the task.

Equally, participants' answers on the OSPAN test were scored strictly and leniently, following distinct procedures. In relation to the strict scoring of the test, one point was credited to each word recalled in the exact order of presentation, that is, in a test set of three trials, for instance, a participant who was able to solve at least two math operations correctly and then recalled their corresponding words, respecting the order of presentation, was given two points. In this scoring procedure, both operation solving and word recall were taken into account. By contrast, the lenient scoring method of the OSPAN test was not as strict in terms of the participants' processing efficiency. In line with previous studies (TURNER; ENGLE, 1989; KANE; BECKLEY; CONWAY; ENGLE, 2001; KANE; CONWAY; HAMBRICK; ENGLE, 2007; PREBIANCA, 2009), a criterion of 85% accuracy in correctly solving all of the mathematical operations of the entire test was required, since it is useful in ensuring that participants do not trade-off between processing the mathematical operations of the test and storing the words that come along with them (UNSWORTH; HEITZ; SCHROCK; ENGLE, 2005). Hence, all words recalled in the exact order of presentation and which obeyed the criterion of 85% accuracy were credited one point, which means that each participant could have gotten up to six wrong mathematical operations (out of a total of 42 trials). If the participant responded to a specific operation of a particular set incorrectly but was able to recall the word following that operation accurately, s/he was credited 1 point as long as s/he had not yet reached six errors.

4 Results and Discussion

4.1 Subtitling, L2 Video Comprehension, and WM Capacity

This section of the paper reports on the results obtained with the correlation tests that were run in order to examine whether participants' WM capacity would present any interaction with their performance on the video comprehension test and/or their performance on the L2 vocabulary test under any experimental (intralingual or interlingual) condition.

To check whether and the extent to which the performance of the participants on the general comprehension part of the video comprehension

test held a significant relationship with their performance on the WM tests, different Spearman's Rank Order Correlation Coefficient tests were run. The results obtained with these tests are reported in Table 1:

TABLE 1 – General Comprehension and WM Correlations

Intralingual Subtitles (n=12)	Correlation Test	GCScore x RST Strict	GCScore x RST Lenient	GCScore x OSPAN Strict	GCScore x OSPAN Lenient
	Spearman's Coefficient r_s		.000	.054	.231
p value		1.000	.867	.470	.286
Interlingual Subtitles (n=12)	Correlation Test	GCScore x RST Strict	GCScore x RST Lenient	GCScore x OSPAN Strict	GCScore x OSPAN Lenient
	Spearman's Coefficient r_s	.556	.517	.148	.302
p value	.061	.085	.645	.340	
Control (n=12)	Correlation Test	GCScore x RST Strict	GCScore x RST Lenient	GCScore x OSPAN Strict	GCScore x OSPAN Lenient
	Spearman's Coefficient r_s	.006	.247	-.177	.297
p value	.986	.439	.581	.348	

Note. n = sample size; RST = Reading Span test; OSPAN = Operation-Word Span Test; p = significance level; GCScore = general comprehension score.

Spearman's Rank Order Correlation Coefficient tests revealed mostly a positive, though not statistically significant correlation between the participants' performance on the general comprehension portion of the video comprehension test and the WM tests at $\alpha = .05$. The correlation tests also revealed two instances of negative correlations, such as in the OSPAN and the general comprehension scores test with the intralingual subtitles group (lenient scoring) and the control group (strict scoring), none of which are statistically significant. The results also indicated a total absence of correlation between the general comprehension scores and the OSPAN Strict variables with the intralingual subtitles groups ($r_s = .000$, $n = 12$, $p > .05$).

To investigate whether and the degree to which the performance of the participants on the specific comprehension part of the video comprehension test would suggest any relationship with their WM capacity, Spearman's Rank Order Correlation Coefficient tests were run. The results obtained are presented in Table 2:

TABLE 2 – Specific Comprehension and WM Correlations

	Correlation Test	SCScore x RST Strict	SCScore x RST Lenient	SCScore x OSPAN Strict	SCScore x OSPAN Lenient
Intralingual Subtitles (n=12)	Spearman's Coefficient r_s	-.338	-.271	-.293	-.520
	p value	.283	.295	.355	.083
	Correlation Test	SCScore x RST Strict	SCScore x RST Lenient	SCScore x OSPAN Strict	SCScore x OSPAN Lenient
	Spearman's Coefficient r_s	.147	.050	.063	.099
Interlingual Subtitles (n=12)	p value	.647	.877	.845	.760
	Correlation Test	SCScore x RST Strict	SCScore x RST Lenient	SCScore x OSPAN Strict	SCScore x OSPAN Lenient
	Spearman's Coefficient r_s	-.146	.000	.271	.167
	p value	.651	1.000	.393	.605
Control (n=12)	Correlation Test	SCScore x RST Strict	SCScore x RST Lenient	SCScore x OSPAN Strict	SCScore x OSPAN Lenient
	Spearman's Coefficient r_s	-.146	.000	.271	.167
	p value	.651	1.000	.393	.605
	Correlation Test	SCScore x RST Strict	SCScore x RST Lenient	SCScore x OSPAN Strict	SCScore x OSPAN Lenient

Note. n = sample size; RST = Reading Span test; OSPAN = Operation-Word Span Test; p = significance level; SCScore = specific comprehension score.

Spearman's Rank Order Correlation Coefficient tests revealed varied results across the groups. In relation to the intralingual subtitles group, the correlations found were negative; as for the interlingual subtitles group, very weak positive correlations were found; finally, regarding the control group, the correlations obtained were either negative or weak positive. None of the correlations found are statistically significant considering the participants' performance on the specific comprehension portion of the video comprehension test and the WM tests ($p > .05$).

As far as the results between the correlation tests performed with the WM measures are concerned – RST and OSPAN tests – and the video comprehension test – general and specific parts – it is clear that, in this study, no statistically significant relationship was found between them in any of the experimental conditions – intralingual and interlingual subtitles – and the control condition. We speculate that some factors might help explain the lack of statistically significant results obtained for those correlational tests.

Firstly, it is important to examine the nature of the WM tests used. When analyzing what participants actually do on the RST, the common understanding is that the test taps one's ability in processing efficiency, mainly as regards comprehension. Daneman and Carpenter (1980), when originally designing the test, argued that individual differences in WM capacity would reflect differences in terms of processing efficiency. They claim that differences in the processing efficiency would be at the core of individual differences regarding language comprehension. In other words, individuals who are not so effective in terms of their processing would possibly have a smaller storage capacity because they would have to free more of their attentional resources to meet the processing demands of a given task. Hence, one's span is likely to correlate with one's capacity to maintain information active in WM long enough to attach it to new information so as to understand a text, for instance. The scholars also claimed that the RST can predict comprehension because it involves many processing aspects that are involved in normal, typical reading.

As for the OSPAN test, Turner and Engle (1989) also intended to measure WM capacity through a task that could tap one's ability to store information for a brief period of time, but that is not language-specific. As demonstrated by Klein and Fiss (1999), the test has been adopted to measure WM capacity because of its high reliability and stability scores. The OSPAN is in tune with Engle and colleagues' view that WM capacity is related to the processing efficiency in any given cognitive task in terms of information processing. Prebianca (2009) explains that “capacity refers to individuals' ability to bring pieces of information from long-term memory into an active state and temporarily maintain that information for further processing by preventing other irrelevant stimuli to enter the focus of attention” (p. 34). Thus, the lack of a significant relationship among the RST and the OSPAN measures and the video comprehension test as well as the L2 vocabulary

recognition tests in the present investigation may be related to the set of relatively different types of tasks employed.

Secondly, while watching the TV series, all three groups had either two or three channels that could compete for their attentional resources. The intralingual subtitles group was submitted to a watching condition in which one auditory channel (soundtrack) and two visual channels (subtitles + video) were simultaneously provided, all of which shared a common language (L2); the interlingual subtitles group, however, was performed in a different condition, as it was provided with both an auditory channel (L2 soundtrack) and two visual channels (L1 subtitles + video), thus being presented with two different linguistic sources of input; as for the control group, two channels were presented to them, that is, one auditory channel (L2 soundtrack) and one visual channel (video).

Participants did not have their attentional resources directed to any of the channels (soundtrack, subtitles, and/or video). In other words, no guiding instructions were provided to them that prompt or influence them in either prioritizing one of the input sources or even attempting to process all of them simultaneously. Hence, participants' attentional resources may have been allocated in only one verbal channel.

Participants in the intralingual subtitles condition outperformed the other two groups on both general and specific comprehension portions of the test (results not included in this paper). Participants performing under that condition were provided with only one language (L2) in all verbal input sources (soundtrack + subtitles). The plausible inference is that there were fewer attention depleting mechanisms involved in comparison to the interlingual subtitles group. This explanation, however, would not be applicable to the control group, whose performance was indeed poorer as compared to the performance of the intralingual subtitles group, but not statistically different from the interlingual subtitles group's performance.

Thirdly, although the RST has been found to correlate mostly with reading comprehension, it may as well be the case that participants' processing on the video comprehension test involved a set of distinct mechanisms, especially taking into consideration the nature of the task employed in this research, involving a multimedia source, and therefore not regular, typical reading. We suspect that one's efficiency in reading subtitles (in L1 or L2) might be more closely linked with one's experience with and the amount of exposure to them. Participants in the present research endeavor

declared being more used to watching subtitled videos with interlingual subtitles than with intralingual subtitles or no subtitles whatsoever.

We also suspected that one's capacity to process the auditory channel of the video would hold a relationship with one's WM capacity – or, more specifically, one's phonological memory – if, in fact, that channel significantly drew on one's attention while processing another channel (subtitles and/or video) simultaneously. In other words, in the present investigation, participants might not have consciously attended to the auditory verbal channel. Given that participants informally stated that they tend to resort to the subtitles more often than the auditory channel, it could be that no seemingly major interaction would arise in this scenario. Selective attention might have played a more prominent role in determining where participants allocated more of their attentional resources to process the input provided in a given channel.

Downing (2000) researched the relationship between selective attention and WM, and contends that selective attention “reduces the load on limited-capacity cognitive systems by filtering irrelevant information from the stimulus stream” (p. 467). The author also explains that in a typical scenario containing many objects, the amount of information present exceeds the capacity of object representation systems, that is, one's capacity to process that (visual) information. Consequently, the objects can be described as “competing” for attention, and the strongest competitors – perhaps the more relevant or salient ones – will be likely to become the focus of selective attention. As a result, they gain access to awareness and guidance of action.

When watching subtitled audiovisual material, viewers possibly focus on the most salient channel, which could be their preferred one or the one that might seem easier for them to process or even one that is more attention drawing. In other words, visual channels – either verbal or non-verbal – could constitute their strongest competitors for attention. The problem in this scenario is that attending to the two channels would require, we assume, a considerable processing load because of the speed in which they are processed (short duration on screen as regards the subtitles and the scenes) and the fact that reading is involved in processing one of the channels. If true, this would explain why participants informally reported that they tend to ignore the soundtrack of the video most of the times.

What Downing (2000) has found in his experiments is that “visual working memory and selective attention share a key functional component: The contents of working memory guide attention even when there is no explicit search task” (p. 469). In the present investigation, participants were not provided with any specific guidance that could direct the allocation of their attentional resources. It is then plausible to assume that when watching subtitled videos, selective attention was directed mostly towards the subtitles and the images of the video.

A somewhat similar perception about this issue is shared by Sydorenko (2010). Her assumptions relate to learners’ efforts to simultaneously allocate attention onto all three modalities. She believes that if learners did pay attention to the three channels simultaneously, this could result in a cognitive overload, which “occurs even when tasks are performed in the native language and is attributed to the limits of working memory” (p. 52).

The contradiction surfaces when L1 and L2 studies are contrasted. Mayer, Heiser, and Lonn (2001), stemming from a cognitive load theory perspective on multimedia learning, found that L1 speakers of English who saw an animation and listened to a simultaneous narration in their L1 were able to gather more information from the narration than those who also received a third modality/channel to process (intralingual subtitles). Their view is that subtitle availability is distracting when soundtrack is also available because they essentially convey the same information, which is thought to follow a redundancy principle in terms of information processing capacity.

However, the literature on subtitling and L2 development has suggested quite the opposite: Subtitle availability has been found to be associated with better L2 comprehension and L2 development. Nonetheless, the preferred source of input to be attended and processed by viewers has not yet been extensively researched. To date, there have only been a handful of studies that have looked at this specific aspect (VANDERPLANK, 1988; TAYLOR, 2005; SYDORENKO, 2010). In general, what these studies have found is that, at first, subtitles are considered distracting by the L2 viewers; with time, viewers reported that they get used to them and even start developing strategies to try to attend to all channels simultaneously.

Fourthly, another possible explanation for the lack of significant correlation between WM and video comprehension could be due to the

sample size and the distribution of high spanners.⁶ Traditionally, research on the relationship between WM and language development has been carried out with much larger sample sizes. In the present investigation, only 12 participants remained until the very end of the data collection in each of the three groups, and the high spanners were grouped in only one experimental group, coincidentally. Therefore, it is possible that no statistically significant correlations were found in this study because of the limited number of participants and the fact that there were fewer high spanners, who were not evenly distributed across experimental and control groups.

4.2 Subtitling, L2 Vocabulary, and WM Capacity

The results reported in this section are an attempt to investigate a possible relationship between the participants' performance on the L2 vocabulary test (including the pre-test, the test, and the post-test) and their WM tests (both the RST and the OSPAN tests). The results obtained with the correlation tests of WM (RST and OSPAN tests) and the L2 vocabulary pre-test are shown first (Table 3), followed by the L2 vocabulary test (Table 4), and finally the L2 vocabulary post-test (Table 5). The discussion of all correlation tests is presented altogether, at the end of this section.

⁶ In this study, as far as the scores on the RST are concerned, three high spanners were found in the data, all of which belong to the interlingual subtitles group; as regard the scores on the OSPAN, three high spanners were also found in the data, all of which belong to the interlingual subtitles group as well.

TABLE 3 – L2 vocabulary pre-test and WM correlations

Intralingual Subtitles (n=12)	Correlation Test	Pre-Test x RST Strict	Pre-Test x RST Lenient	Pre-Test x OSPAN Strict	Pre-Test x OSPAN Lenient
	Spearman's Coefficient r_s	-.419	-.440	-.262	-.102
	<i>p</i> value	.175	.152	.411	.753
Interlingual Subtitles (n=12)	Correlation Test	Pre-Test x RST Strict	Pre-Test x RST Lenient	Pre-Test OSPAN Strict	Pre-Test x OSPAN Lenient
	Spearman's Coefficient r_s	-.194	-.112	.178	.130
	<i>p</i> value	.546	.729	.580	.687
Control (n=12)	Correlation Test	Pre-Test x RST Strict	Pre-Test x RST Lenient	Pre-Test x OSPAN Strict	Pre-Test x OSPAN Lenient
	Spearman's Coefficient r_s	.169	.264	-.184	.286
	<i>p</i> value	.599	.407	.568	.367

Note. n = sample size; RST = Reading Span test; OSPAN = Operation-Word Span Test;
p = significance level.

TABLE 4 – L2 vocabulary test and WM correlations

	Correlation Test	Test x	Test x	Test x	Test x
		RST Strict	RST Lenient	OSPAN Strict	OSPAN Lenient
Intralingual Subtitles (n=12)	Spearman's Coefficient r_s	-.119	-.254	-.209	-.100
	p value	.713	.426	.515	.756
	Correlation Test	Test x RST Strict	Test x RST Lenient	Test x OSPAN Strict	Test x OSPAN Lenient
Interlingual Subtitles (n=12)	Spearman's Coefficient r_s	-.323	-.281	.360	.230
	p value	.305	.377	.251	.473
	Correlation Test	Test x RST Strict	Test x RST Lenient	Test x OSPAN Strict	Test x OSPAN Lenient
Control (n=12)	Spearman's Coefficient r_s	.203	.318	-.029	.405
	p value	.526	.314	.928	.192
	Correlation Test	Test x RST Strict	Test x RST Lenient	Test x OSPAN Strict	Test x OSPAN Lenient

Note. n = sample size; RST = Reading Span test; OSPAN = Operation-Word Span Test;
 p = significance level.

TABLE 5 – L2 vocabulary post-test and WM correlations

Intralingual Subtitles (n=12)	Correlation Test	Post-Test x RST Strict	Post-Test x RST Lenient	Post-Test x OSPAN Strict	Post-Test x OSPAN Lenient
	Spearman's Coefficient r_s	.152	-.022	-.283	-.391
<i>p</i> value	.638	.945	.373	.209	
Interlingual Subtitles (n=12)	Correlation Test	Post-Test x RST Strict	Post-Test x RST Lenient	Post-Test x OSPAN Strict	Post-Test x OSPAN Lenient
	Spearman's Coefficient r_s	-.256	-.235	.268	.081
<i>p</i> value	.421	.462	.399	.803	
Control (n=12)	Correlation Test	Post-Test x RST Strict	Post-Test x RST Lenient	Post-Test x OSPAN Strict	Post-Test x OSPAN Lenient
	Spearman's Coefficient r_s	.364	.487	.051	.276
<i>p</i> value	.245	.109	.876	.385	

Note. n = sample size; RST = Reading Span test; OSPAN = Operation-Word Span Test; *p* = significance level.

As shown in Table 3, the results obtained with the tests revealed varied results for the three groups, converging into mostly moderate, weak negative correlations and a few weak positive correlations, all of which are not statistically significant ($p > .05$). In other words, the results of the correlation tests are suggestive of an absence of significant relationship between the participants' performance on the L2 vocabulary pre-test and the WM tests. Moreover, as indicated in Table 4, the results obtained with correlation tests revealed varied results across the groups, converging into mostly weak negative correlations, though there are a few positive weak and moderate correlations. No statistically significant relationship was found ($p > .05$), which can be interpreted as an absence of noteworthy association between the participants' performance on the L2 vocabulary test and the WM tests. Finally, not surprisingly, the results reported in Table 5 also

revealed varied results across the groups, which range from an almost perfect absence of correlation to either a moderate negative or positive correlation. However, all of correlations obtained are not statistically significant ($p > .05$), which are indicative of an absence of substantial relationship between the participants' performance on the L2 vocabulary post-test and on the WM tests in the data analyzed herein.

Firstly, provided that WM measures, especially the OSPAN test, have mostly been found to correlate with lexical knowledge development, it would be realistic and possible to find some significant interaction between participants' WM measures and their L2 vocabulary test results, mostly in the immediate L2 vocabulary test. This expectation in part stems from the fact that no treatment was provided in the pre-test; as for the post-test, a one-week delay was adopted, so results could possibly interfere with long-term memory.

One of the reasons underlying the lack of substantial correlations among the WM measures and the L2 vocabulary recognition tests might be associated with the nature of the tests regarding time and storage. On both RST and OSPAN tests, participants had to hold some piece of information (words) for a very brief period of time in their WM, maintaining it active, and then had to recall that very same piece of information to verbalize it (saying it out loud when question marks appeared on screen). Unlike the WM tests, while watching the video in any of the experimental or control conditions, participants might not have been consciously trying to hold any piece of information active for later retrieval. Hence, the nature of the tasks involved – WM and L2 vocabulary recognition – is apparently dissimilar, particularly in terms of the processing demands required to accomplish each of them.

Secondly, as Baddeley (2009) explains, short-term memory (STM) and the system(s) responsible for it are part of WM. The author also explains that STM involves the capacity to store small amounts of information for very brief periods of time. Moreover, Schwartz and Metcalfe (1992) explain that the STM is deemed an active memory with limited capacity to hold information for around twenty to sixty seconds. The information received in the STM is assumed to be stored for a short period of time while being analyzed and interpreted. Once understood, part of it is transmitted to the LTM, possibly for permanent storage. The old information that is no longer needed may fade away from the STM (MAYER; MORENO, 1998).

Thus, the operations going on in the STM are thought to be indispensable for long-lasting storage.

In order for information to be stored in the STM and later transmitted to the LTM, it must be analyzed and possibly understood, otherwise it will be lost or discarded. However, information will only be analyzed if such a piece of information is, we presume, attended to. This view that links the STM to consciousness is shared by a few scholars (BAARS, 1986; SCHMIDT, 1990). In sum, their perception is that the STM serves as a type of “broadcasting station” (BAARS, 1986 *apud* SCHMIDT, 1990, p. 137) and stresses the role of consciousness to actuate the learning process. Although consciousness in learning has been controversially debated in the field of Second Language Acquisition, many cognitive psychologists, such as Rutherford and Smith (1985); Schmidt (1990); Koskinen and colleagues (1996); and Hsiao and Oxford (2002), mostly agree upon the need to raise learners’ consciousness to raise their awareness of the properties of what to learn.

In the L2 vocabulary test, more specifically the immediate test, participants watched a 20-minute video and were then given the L2 vocabulary task. It is possible to hypothesize that participants noticed some of the target lexical items, understood them, and inferred their meaning or translation counterparts, but by the time they were administered the test – that is, 20 minutes later – some of those items may have no longer been active in their STM for some reason, which could be dependent on a larger storage and processing capacity on their part. Moreover, the simple fact that participants may have been more concerned with understanding the story and enjoying the comedy would suggest that their attentional resources were not consciously allocated towards the lexical items.

Some of these lexical items, unfortunately, may not have even been integrated into their STM in the first place, because some of them may not have been processed, understood, and therefore registered. Furthermore, it is also plausible that lexical items were not held active in participants’ STM because of a decay in a memory trace due to a lack of frequency/saliency in the input. Following a similar line of reasoning, it might be the case that, because very few of the participants were high spanners, in general, their capacity to hold active pieces of information about the video may have proven insufficient for later retrieval.

Thirdly, the redundancy produced by subtitled materials may present some degree of influence upon consciousness and STM capacity. These theorists argue that visual channels (subtitles and video) need a certain amount of attention for processing because of their dynamic and graphic nature. They claim that the pictures that require attention can be processed with little effort. Consequently, more attention resources or capacity should be available to be used (D'YDEWALLE *et al.*, 1991; GRIMES, 1991). If true, this assumption would mean that if one understands the story narrated on screen more easily, more attentional resources should be available to focus on other aspects of the subtitled video. Looking at the results of the video comprehension test and L2 vocabulary tests, it is clear that the intralingual subtitles group outperformed the other two groups. Even though these performances did not always achieve statistical significance in the comparison across groups, the numbers could suggest that the availability of subtitles – and mostly English subtitles – enabled better video comprehension and might have made more attentional resources available to those participants performing in either experimental conditions to examine the linguistic content in the input.

It is possible that, from a statistical point of view, no significant interaction was found between WM measures and L2 vocabulary because of the low scores obtained on the L2 vocabulary recognition tests (not reported here due to space constraints), as well as the limited sample size in each experimental and control group. It is also possible that no statistically significant interaction between WM measures and L2 vocabulary recognition tests was found because of the limited number of high spanners. Finally, it is reasonable that the WM tests and other measures used in the present study were not able to capture the complexities involved in the actual processing of subtitled videos, and therefore significant correlations may not have surfaced due to the nature of the tests adopted in this study.

5 Final Remarks

This mixed design study, focusing on qualitative and quantitative analyses, set out to investigate whether or not L2 learners' WM capacity could show any relationship with their performance on subtitled video materials' processing tasks, namely an L2 video comprehension test and an L2 vocabulary test. Though no statistically significant results were found,

we need to bear in mind that correlation analyses cannot be interpreted as establishing cause-and-effect relationships. In our understanding, correlation analyses allow, with some degree of caution, to indicate the level of association among variables. Therefore, the results obtained herein demonstrate that the variables investigated did not interact significantly. It may also be the case that WM could possibly be related to participants' performance on the tests they took, to a lesser or greater degree. Different results might be obtained with larger sample sizes and modifications in the instruments or procedures for data collection.

Further research is still necessary to deepen our current understanding of the impact that subtitled video materials can have on L2 learners. Furthermore, given the expanding use of subtitled audiovisual materials in and outside L2 classrooms and the technological advances we are currently witnessing, we find it crucial for L2 practitioners to help learners develop strategies to benefit from subtitled video materials, especially in relation to less strategic L2 learners. Finally, L2 instructors should consider helping learners to cope with the possible cognitive (over)load that these materials may present, most of all when any textual aid (subtitles) is used. Directing learners' attention to specific features of the video material presented might be a good start to achieve this end.

Acknowledgment

This study was sponsored by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – Brazil, granted to the first author of this paper.

References

- BAARS, B. *The cognitive revolution in psychology*. New York: Guilford, 1986.
- BADDELEY, A. D. Working memory. *Science*, Washington, v. 255, p. 556-559, 1992. Doi: <https://doi.org/10.1126/science.1736359>
- BADDELEY, A. D. The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, [s.l.], v. 4, n. 11 p. 417-423, 2000.
- BADDELEY, A. D. Short-term Memory. In: BADDELEY, A. D.; EYSENCK, M.; ANDERSON, M. C. (Org.) *Memory*. Hove: Psychology Press, 2009.

- BADDELEY, A. D. Estimating the short-term component in free recall. *British Journal of Psychology*, [s.l.], v. 61, n. 1, p. 13-15, 2011.
- BIANCHI, F.; CIABATTONI, T. *Captions and subtitles in EFL learning an investigative study in a comprehensive computer environment*. Trieste: Edizioni Università di Trieste, 2008.
- BIRD, S.; WILLIAMS, J. The effect of bimodal input on implicit and explicit memory: an investigation into the benefits of within-language subtitling. *Applied Psycholinguistics*, Cambridge, v. 23, n. 4, p. 509-533, 2002. Doi: <https://doi.org/10.1017/S0142716402004022>
- CAIMI, A. Audiovisual Translation and Language Learning. The Promotion of Intralingual Subtitles. *Journal of Specialised Translation*, London, v. 6, n. 1, p. 85-98, 2006.
- CHANG, S. The Interaction between Schemata and Subtitles. *Journal of National Taipei University of Technology*, Taipei, v. 39, n 1, p. 209-227, 2006.
- CONWAY, A R.; KANE, M. J.; BUNTING, M. F.; HAMBRICK, D. Z.; WILHELM, O.; ENGLE, R. W. Working memory span tasks: *A methodological review and user's guide*. *Psychonomic Bulletin & Review*, [s.l.], v. 12, n. 5, p. 769-786, 2005.
- DANAN, M. Captioning and subtitling: Undervalued language learning strategies. *Meta*, Montreal, v. 49, n. 1, p. 67-77, 2004. Doi: <https://doi.org/10.7202/009021ar>
- DANEMAN, M.; CARPENTER, P. A. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, [s.l.], v. 19, n. 4 p. 450-466, 1980.
- D'YDEWALLE, F.; GIELEN, I. Attention allocation with overlapping sound, image, and text. In: RAYNER, K. (Org.). *Eye movements and visual cognition: Scene perception and reading*. New York: Springer-Verlag, 1992. p. 415-427. Doi: https://doi.org/10.1007/978-1-4612-2852-3_25
- D'YDEWALLE, G.; PRAET, C.; VERFAILLIE, K.; VAN RENSBERGEN, J. Watching subtitled television: Automatic reading behavior. *Communication Research*, Santa Barbara, v. 18, n. 5, p. 650-666, 1991. Doi: <https://doi.org/10.1177/009365091018005005>
- D'YDEWALLE, G.; VAN DE POEL, M. Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research*, [s.l.], v. 28, n. 3, p. 227-244, 1999.
- DOWNING, P. E. Interactions between visual working memory and selective attention. *Psychological Science*, [s.l.], v. 11, n. 6, p. 467-473, 2000.

GRIMES, T. Mild auditory-visual dissonance in television news may exceed viewer attentional capacity. *Human Communication Research*, Seoul, v. 18, n. 2, p. 268-298, 1991. Doi: <https://doi.org/10.1111/j.1468-2958.1991.tb00546.x>

HAYATI, A.; MOHMEDI, F. The effect of films with and without subtitles on listening comprehension of ESL learners. *British Journal of Educational Technology*, Cardiff, v. 42, n. 1, p.181-192, 2011. Doi: <https://doi.org/10.1111/j.1467-8535.2009.01004.x>

HSIAO, T. Y.; OXFORD, R. Comparing Theories of Language Learning Strategies: A confirmatory Factor Analysis. *The Modern Language Journal*, Indianapolis, v. 86, n. 3, p. 368-383, 2002. Doi: <https://doi.org/10.1111/1540-4781.00155>

HUANG, H.; ESKEY, D. E. The effects of closed-captioned television on the listening comprehension of intermediate English as a second language (ESL) students. *Journal of Educational Technology Systems*, New York, v. 28, n. 1, p. 75-96, 1999. Doi: <https://doi.org/10.2190/RG06-LYWB-216Y-R27G>

KANE, M. J.; BECKLEY, M. K.; CONWAY, A. R.; ENGLE, R. W. A controlled-attention view of working memory capacity. *Journal of Experimental Psychology: General*, [s.l.], v. 130, n. 2, p. 169-183, 2001.

KANE, M. J.; CONWAY, A. R. A.; HAMBRICK, D. Z.; ENGLE, R. W. Variation in working-memory capacity as variation in executive attention and control. In: CONWAY, A. R. A.; JARROLD, C.; KANE, M. J. A.; MIYAKE, J. T. (Org). *Variation in Working Memory*. New York: Oxford University Press, 2007.

KLEIN, K.; FISS, W. H. The reliability and stability of the Turner and Engle working memory task. *Behavior Research Methods, Instruments, & Computers*, [s.l.], v. 31, n. 3, p. 429-432, 1999.

KOOSLTRA, C. M.; BEENTJES, J. W. J. Children vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development*, [s.l.], v. 47, n. 1, p. 51-60, 1999.

KOSKINEN, P. S.; KNABLE, J. E.; MARKHAM, P. L.; JENSEMA, C. J.; KANE, K. W. Captioned Television and the Vocabulary Acquisition of Adult Second Language Correctional Facility Residents. *Journal of Educational Technology Systems*, New York, v. 24, n. 4, p. 359-373, 1996. Doi: <https://doi.org/10.2190/JKC2-LE5D-36BC-809F>

KRUGER, J.; STEYN, F. Subtitles and Eye Tracking: Reading and Performance. *Reading Research Quarterly*, Nashville, v. 49, n. 1, p. 105-120, 2014. Doi: <https://doi.org/10.1002/rrq.59>

LATIFI, M.; MOBALEGH, A.; MOHAMMADI, E. Movie Subtitles and the Improvement of Listening Comprehension Ability. Does it help? *Journal of Language Teaching and Learning*, [s.l.], v. 1, n. 2, p. 18-29, 2011.

LAUFER, B. What's in a word that makes it hard or easy? Intralexical factors affecting the difficulty of vocabulary acquisition. In: MCCARTHY, M.; SCHMITT, N. (Org.). *Vocabulary Description, Acquisition and Pedagogy*. New York: Cambridge University Press, 1997. p. 140-155.

MARKHAM, P. L. Captioned videotapes and second-language listening word recognition. *Foreign Language Annals*, Michigan, v. 32, n. 3, p. 321-328, 1999. Doi: <https://doi.org/10.1111/j.1944-9720.1999.tb01344.x>

MARKHAM, P. L.; PETER, L. The influence of English language and Spanish language captions on foreign language listening/reading comprehension. *Journal of Educational Technology Systems*, New York, v. 31, n. 3, p. 331-341, 2003. Doi: <https://doi.org/10.2190/BHUH-420B-FE23-ALA0>

MARKHAM, P. L.; PETER, L.; MCCARTHY, T. The effects of native language vs. target language captions on foreign language students' DVD video comprehension. *Foreign Language Annals*, Michigan, v. 34, n. 5, p. 439-445, 2001. Doi: <https://doi.org/10.1111/j.1944-9720.2001.tb02083.x>

MATIELO, R. Intralingual Subtitles, Interlingual Subtitles, and L2 Vocabulary Learning: An Exploratory Study with Brazilian EFL Students. 2016. 272 f. Dissertation (Ph.D. in English and Applied Linguistics) – Universidade Federal de Santa Catarina, Florianópolis, 2016.

MATIELO, R.; COLLET, T.; D'ELY, R. C. S. F. The effects of interlingual and intralingual subtitles on vocabulary learning by Brazilian EFL learners. An exploratory study. *Revista Intercâmbio*, Montes Claros, v. 27, n. 1, p. 83-99, 2013.

MATIELO, R.; D'ELY, R. C. S. F.; BARETTA, L. The effects of interlingual and intralingual subtitles on second language learning/acquisition: a state-of-the-art review. *Trabalhos em Linguística Aplicada*, Campinas, v. 54, n. 1, p. 161-182, 2015. Doi: <https://doi.org/10.1590/0103-18134456147091>

MATIELO, R.; OLIVEIRA, R. P.; BARETTA, L. Intralingual Subtitles, Interlingual Subtitles, and Video Comprehension: Insights from an Exploratory Study. *Letrônica*, Porto Alegre, v. 10, n. 2, p. 758-774, 2017. Doi: <https://doi.org/10.15448/1984-4301.2017.2.26370>

MATIELO, R.; OLIVEIRA, R. P.; BARETTA, L. Intralingual Subtitles, Interlingual Subtitles, and L2 Vocabulary: Developments from an Exploratory Study. *Acta Scientiarum*, Maringá. Forthcoming. Doi: <https://doi.org/10.4025/actascilangcult.v40i1.36570>

MAYER, R.; MORENO, R. A Split-Attention Effect in Multimedia Learning: Evidence for Dual Processing Systems in Working Memory. *Journal of Educational Psychology*, Washington, v. 90, n. 2, p. 312-320, 1998. Doi: <https://doi.org/10.1037/0022-0663.90.2.312>

MAYER, R. E.; HEISER, J.; LONN, S. Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, Washington, v. 93, n. 1, p. 187-198, 2001. Doi: <https://doi.org/10.1037/0022-0663.93.1.187>

MONTERO-PEREZ, M; PETERS, E; DESMET, P. Enhancing Vocabulary Learning through Captioned Video. An Eye-Tracking Study. *Modern Language Journal*, Indianapolis, v. 99, n. 2, p. 308-328, 2015. Doi: <https://doi.org/10.1111/modl.12215>

PEREGO, E.; DEL MISSIER, F.; PORTA, M.; MOSCONI, M. The Cognitive Effectiveness of Subtitle Processing. *Media Psychology*, [s.l.], v. 13, n. 3, p. 243-272, 2010.

PREBIANCA, G. V. V. Working memory capacity, lexical access and proficiency level in L2 speech production. Dissertation (Ph.D. in English and Applied Linguistics) – Post-Graduation in Letters/English and Corresponding Literature, Universidade Federal de Santa Catarina – Florianópolis, 2009.

PRICE, K. Closed-Captioned TV. An Untapped Resource. *MATSOL Newsletter*, [s.l.], v. 12, n. 2, p. 1-8, 1983.

RAINE, P. Incidental Increase in Depth of Vocabulary Knowledge Through the Viewing of Subtitled, Authentic Videos. In: SONDA, N.; KRAUSE, A. (Org.). *Making a Difference*. Tokyo: JALT, 2013. p. 492-505.

RUTHERFORD, W.; SMITH, M. S. Consciousness-raising and universal grammar. *Applied Linguistics*, Cambridge, v. 6, n. 3, p. 274-82, 1985. Doi: <https://doi.org/10.1093/applin/6.3.274>

SCHMIDT, R. The Role of Consciousness in Second Language Learning. *Applied Linguistics*, Cambridge, v. 11, n. 2, p. 129-158, 1990. Doi: <https://doi.org/10.1093/applin/11.2.129>

SHARIF, M. R.; EBRAHIMIAN, M. A Comparative Study of ‘Subtitled’ vs. ‘Auditory’ Documentaries Comprehension among Persian Speakers of English.

International Journal of Applied Linguistics & English Literature, Footscray, v. 2, n. 2, p.185-192, 2013. Doi: <https://doi.org/10.7575/aiac.ijalel.v2n.2p.185>

STEWART, M. A.; PERTUSA, I. Gains to language learners from viewing target language closed-captioned films. *Foreign Language Annals*, Michigan, v. 37, n. 3, p. 438-447, 2004. Doi: <https://doi.org/10.1111/j.1944-9720.2004.tb02701.x>

SYDORENKO, T. Modality of Input and Vocabulary Acquisition. *Language Learning & Technology*, Mānoa, v. 14, n. 2, p. 50-73, 2010.

TAYLOR, G. Perceived processing strategies of students watching captioned video. *Foreign Language Annals*, Michigan, v. 38, n. 3, p. 422-427, 2005. Doi: <https://doi.org/10.1111/j.1944-9720.2005.tb02228.x>

TORRES, A. C. G. Working memory capacity and reader's performance on main idea construction in L1 and L2. Dissertation (Ph.D. in English and Applied Linguistics) – Post-graduation in Letters/English and Corresponding Literature, Universidade Federal de Santa Catarina – Florianópolis, 2003.

TURNER, M. L.; ENGLE, R. W. Is working memory capacity task dependent? *Journal of Memory and Language*, New York, v. 28, n. 2, p. 127-154, 1989. Doi: [https://doi.org/10.1016/0749-596X\(89\)90040-5](https://doi.org/10.1016/0749-596X(89)90040-5)

UNSWORTH, N.; HEITZ, R. P.; SCHROCK, J. C.; ENGLE, R. W. An automated version of the operation span task. *Behavior Research Methods*, [s.l.], v. 37, n. 3, p. 498-505, 2005.

VAN LOMMEL, S.; LAENEN, A.; D'YDEWALLE, G. Foreign grammar acquisition while watching subtitled television programs. *British Journal of Educational Psychology*, London, v. 76, n. 1, p. 243-258, 2006. Doi: <https://doi.org/10.1348/000709905X38946>

VANDERPLANK, R. The value of teletext subtitles in language learning. *ELT Journal*, Oxford, v. 42, n. 4, p. 272-281, 1988. Doi: <https://doi.org/10.1093/elt/42.4.272>

VANDERPLANK, R. *Captioned Media in Foreign Language Learning and Teaching. Subtitles for the Deaf and Hard-of-Hearing as Tools for Language Learning*. London: Palgrave Macmillan, 2015.

WINKE, P.; GASS, S.; SYDORENKO, T. The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, Mānoa, v. 14, n. 1, p. 65-86, 2010.

WINKE, P.; GASS, S.; SYDORENKO, T. Factors Influencing the Use of Captions by Foreign Language Learners. An Eye-Tracking Study. *Modern Language Journal*,

Indianapolis, v. 97, n. 1, p. 254-275, 2013. Doi: <https://doi.org/10.1111/j.1540-4781.2013.01432.x>

ZAREI, A. A.; GILANIAN, M. The effect of multimedia modes on L2 vocabulary learning. *International Journal of Management and Humanity Sciences*, [s.l.], v. 2, n. 1, p. 1011-1020, 2013.

ZAREI, A. A.; RASHVAND, Z. The Effect of Interlingual and Intralingual, Verbatim and Nonverbatim Subtitles on L2 Vocabulary Comprehension and Production. *Journal of Language Teaching and Research*, [s.l.], v. 2, n. 3, p. 618-625, 2011.

Data de submissão: 29/11/2017. Data de aprovação: 03/05/2018.