

The effects of binge-watching and spacing on learning L2 multi-word units from captioned TV series

Anastasia Pattemore & Carmen Muñoz

To cite this article: Anastasia Pattemore & Carmen Muñoz (2023): The effects of binge-watching and spacing on learning L2 multi-word units from captioned TV series, The Language Learning Journal, DOI: [10.1080/09571736.2023.2211614](https://doi.org/10.1080/09571736.2023.2211614)

To link to this article: <https://doi.org/10.1080/09571736.2023.2211614>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 17 May 2023.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

The effects of binge-watching and spacing on learning L2 multi-word units from captioned TV series

Anastasia Pattermore  ^{a,b,c} and Carmen Muñoz  ^b

^aFaculty of Arts, Department of European Languages and Cultures, University of Groningen, Groningen, Netherlands; ^bDepartment of Modern Languages and Literatures and English Studies, University of Barcelona, Barcelona, Spain; ^cDepartment of Foreign Languages and Communication Technologies, NUST MISiS, Moscow, Russia

ABSTRACT

The present study explores the effects of distributed practice by extending this area of research to L2 learning from audiovisual input. A total of 96 L1 Russian elementary to advanced learners of English watched five episodes of captioned TV series under three viewing distribution conditions: longer spacing of viewing once per week; shorter spacing of viewing once per day; and massed binge-watching of viewing all the episodes in one session. The participants' knowledge of 25 target multi-word units was tested via the Vocabulary Knowledge Test before and after the intervention. The analyses considered learners' viewing conditions, proficiency level and feeling of learning as factors. The results indicated that the participants significantly improved their knowledge of the target multi-word units regardless of their viewing condition, but there was a benefit of longer spacing compared to the other two groups. Although binge-watching is an enjoyable everyday activity for many viewers, it might be advisable for L2 learners to space their viewing sessions to achieve better outcomes.

ARTICLE HISTORY


Received 13 March 2023
Accepted 3 May 2023

KEYWORDS

Audiovisual input; captions; time distribution; binge-watching; spacing; multi-word units

Introduction

Audiovisual or multimodal materials are recognised by language learners and teachers as an excellent way to receive exposure to the target language. Specifically, original version (OV) TV series have the potential to provide learners with large amounts of spoken input (Webb 2014). Furthermore, it has been claimed that the addition of captions – a simultaneous, on-screen written text representation of the soundtrack – benefits language learners because the three input sources (audio, visual and caption text) complement each other and support learning from audiovisual input by distributing the information among the three channels (see Vanderplank 2016). Captioning has proven to be a useful technique to support L2 listening comprehension, vocabulary, grammar, collocations and pronunciation (see Montero Perez 2022). Previous audiovisual studies have focused on factors such as individual differences, and type of on-screen text. However, little is known about the role of time distribution – how much time should be left between viewings, for example between consecutive episodes of a TV series. While there have been studies on time distribution in reading-while-listening (e.g. Serrano and Huang 2018) audiovisual input has only started to be examined from this perspective (Muñoz et al. 2022). In particular, the phenomenon of binge-watching – watching at

CONTACT Anastasia Pattermore  a.pattermore@rug.nl

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

least two TV series episodes consecutively (Castro et al. 2021) – has not been explored from the second language acquisition (SLA) point of view.

As binge-watching has become more extended during the times of COVID-19 lockdown and is becoming an ‘established viewing norm’ (Rubenking and Bracken 2021), it is likely that this increase is not only limited to L1 viewers. Several studies point out that watching L2 TV is one of the most popular methods of engaging with a foreign language outside of the classroom (De Wilde et al. 2021; Lindgren and Muñoz, 2013; Muñoz 2020). Additionally, with the increasing use of streaming platforms, it is probable that the growth of binge-watching will also be seen in L2 media consumption.

However, some behavioural researchers are concerned about binge-watching as a viewing pattern. For instance, Steins-Loebers et al. (2020) suggest negative consequences of binge-watching (e.g. fatigue). From the cognitive psychology perspective, binge-watching could be associated with massed exposure to input¹, which may be less effective than spacing the learning sessions (Carpenter 2017).

The present study explores the effects of three different viewing time distributions of Original Version (OV) TV series on learning of multi-word units (MWUs) and addresses the role of L2 proficiency level and feeling of learning from audiovisual input.

Literature review

Learning from audiovisual input

Recent research has been consistent regarding the positive effects of OV television for L2 development (see Montero Perez 2022). However, mere extensive exposure to audiovisual input is unlikely to be sufficient in itself for language learning to occur. It has been found, for instance, that to process uncaptioned audiovisual input efficiently, L2 learners need to have high levels of working memory capacity (Pattemore and Muñoz 2020). On the other hand, the addition of captions attenuates the limits in the working memory capacity thus providing a necessary scaffold. A few studies have provided robust evidence of an advantage of viewing OV media with captions rather than without, for vocabulary, grammar and speech processing (Montero Perez 2022). This advantage could be explained through input processing theories such as the ‘subtitle principle’ (Mayer et al. 2020) that proposes that on-screen text provides L2 viewers with an opportunity to revisit the language content if they were unable to fully process the audio. This way, viewers would have a ‘backup option’ and their working memory would not become overloaded. Rather than overloading the viewer’s cognitive capacity, captions would maximise L2 learning from audiovisual input, as the three input sources balance each other’s processing loads (Frumuselu et al. 2015).

Along with cognitive individual differences, another factor that may explain the rate of learning from audiovisual input, and one of the variables in this study, is the viewer’s proficiency level. Individuals’ L2 proficiency influence on learning from audiovisual material has been frequently discussed (e.g. Vanderplank 2016: 80). It has been suggested that an intermediate level is the necessary proficiency threshold to start benefiting from captioned audiovisual input to the full extent (e.g. Danan 2004). In addition, vocabulary research suggests that the Matthew effect of ‘the rich get richer’ comes into play where viewers with larger vocabulary size have higher learning gains (Suárez and Gesa 2019). In grammar studies, on the other hand, it seems that intermediate to upper-intermediate proficiency levels benefit the most from extensive viewing of TV series. It has been suggested that the lack of challenging grammatical constructions in the input may hinder advanced learners’ progress (Pattemore and Muñoz 2020).

Though most studies in this area have focused on vocabulary, only a few have explored learning of MWUs, that is, ‘units longer than a single word’ (Majuddin et al. 2021), also referred to as formulaic sequences (Pui-même and Peters 2019), and fully filled constructions (Pattemore and Muñoz 2022). A recent study on the learning of multi-word expressions through OV TV series included a comparison between no captions and captions (Majuddin et al. 2021). The participants were exposed to a single (20 minutes) or repeated viewing of one episode of a TV series. There was no group difference in the

repeated viewing condition, while the results of the immediate post-test for the single viewing condition yielded a significant effect of captioning over no captions.

The results of the abovementioned study suggest that MWUs can be learnt from audiovisual input, with an advantage of watching with L2 captions, but no previous research has explored the effect of viewing time distribution on learning of MWUs.

Distributed practice effect in psychology research

It is well accepted among psychology researchers that learning is more robust when study takes place in a spaced, distributed manner (Carpenter 2017; Küpper-Tetzl 2014; Toppino and Gerbier 2014). The spacing effect is concerned with the effects of spacing and massing, and it is a very strong effect in cognitive psychology. It suggests that items that appear in spaced sequences (with some time between encounters) are better learned and remembered than items in massed sequences (immediate repetitions). Carpenter (2017) explained this advantage of spacing by several mechanisms, such as retrieval effort, encoding variability and deficient processing. Retrieval effort is defined as the process of recalling previously learned information when learners have to turn to their memory trace to retrieve the meaning of the target item. When a new item is encountered in context following longer spaced distribution, each subsequent encounter after the first one represents an opportunity to retrieve stored information about the target item from previous encounters. In contrast, multiple instances of the same new item close together (massed distribution) are likely to be encoded as one learning episode; thus, retrieval from previous learning episodes is unlikely. Close to the retrieval effort is encoding variability. In spaced conditions, each encounter with the target item is associated with specific contextual cues the learners find themselves in (e.g. specific day, time, events and learner's mood). These contextual cues, according to the encoding-variability hypothesis (Martin 1968), increase the possibility of successful retrieval.

In the massed condition, on the other hand, this type of variability in associations with learning may not be present, as it takes place during the same session. Finally, when the learning episodes happen in a massed manner, multiple exposures to target items become redundant, causing decline in attentional processing and, consequently, lower learning gains compared to spaced exposure (Carpenter 2017).

When it comes to comparing different types of spacing, it has been suggested that the time lag between the spaced sessions affects the learning outcomes. The lag effect concept pertains to the effects of different amounts of spacing between the learning sessions (e.g. one day or one week). In general, the lag effect is a phenomenon where longer spacing mostly leads to better long-term retention than shorter spacing (Cepeda et al. 2008). It has been suggested that retrieval effort is stronger in the longer spacing condition as it implies a more challenging task of recalling previously encountered information (Carpenter 2017). In the same line, while spacing is related to better consolidation, if too little time has passed between the learning sessions, the consolidation might not have taken place and the effects of spacing could be limited. On the other hand, if the spacing between the sessions is too long, previous presentations might fail to be retrieved, and in consequence, the advantage of longer spacing may disappear (Carpenter 2017).

Although the results of objective tests demonstrate an advantage of spacing over massing, the participants might have different ideas about their learning process, as in several studies learners reported greater feeling of learning from massed than from spaced exposure (e.g. Birnbaum et al. 2013). These beliefs might affect learner's decisions about massing or spacing their individual learning sessions (Carpenter 2017).

Distributed practice in SLA

While educational and cognitive psychology research suggests that learning in a spaced manner leads to better outcomes, when it comes to L2 learning, the spacing advantage is not so clear

(Kim and Webb 2022; Serrano 2012). For example, studies that explored the distribution of target items' occurrence (e.g. in how many book chapters or individual books the items appeared) have had mixed results. Webb and Chang (2015) analysed the effects of distribution of occurrence of target vocabulary in reading-while-listening of ten graded readers. The results suggested no significant correlation between the items' learning gains and the distribution of occurrence. Contrastingly, Elgort and Warren (2014) in their study of learning pseudowords from reading several chapters of an authentic English book, found a benefit of having items massed in the same chapter. The massing advantage was observed especially for the high-intermediate rather than advanced L2 learners, as the high-intermediate group's learning dropped significantly when the items were distributed across book chapters.

Serrano and Huang (2018) explored the distribution of reading-while-listening sessions on incidental learning of vocabulary from repeated reading. The participants were divided into two groups with different inter-session intervals (one day and seven days between the sessions). While there was an advantage of shorter spacing at the immediate post-test, a benefit of longer spacing appeared in the delayed post-test, suggesting that longer spacing leads to better retention. For intentional learning, on the other hand, Serrano and Huang (2021) found an advantage of shorter spacing for both short-term and long-term vocabulary retention. Further, Elgort et al. (2018) proposed that the encounters with the target words across multiple sessions are more beneficial for the learning of meaning, while the massed condition where learners encounter the items multiple times in a single session, leads to more successful word form learning. Considering L2 grammar learning, distributed practice may affect different types of knowledge. For instance, long spacing may be more effective for the acquisition of explicit knowledge (Bird 2010), while short spacing could be beneficial for knowledge proceduralisation and automatisisation (Suzuki and DeKeyser 2017).

In their recent meta-analysis, Kim and Webb (2022) analysed the effect sizes from 48 experiments on spaced practice in SLA and focused on moderating variables such as number of sessions and target items. The results showed that spacing had a medium to large effect on L2 learning, and that spaced practice was significantly more effective than massed practice, supporting the findings in cognitive psychology research (Carpenter 2017). Considering the lag effect, there was no difference between the shorter and longer spacing in immediate posttests, but longer spacing was more effective based on delayed posttest results. Interestingly, the benefit of longer spacing was only observed when the experiments involved a sole learning session. When the studies consisted of multiple learning sessions, the lag effect diminished. As for the language features, spacing promoted better results for learning vocabulary than grammar. Longer, rather than shorter, spacing was more beneficial for L2 vocabulary. Finally, shorter spacing was more effective for L2 grammar immediate learning, but longer spacing encouraged better retention.

As can be seen from the literature reviewed above, there is no consensus about the benefits of distributed practice in SLA. Importantly, conclusions from previous research suggest that the distribution of items or sessions is not the only factor at play, as such factors as learners' language proficiency (Elgort and Warren 2014), and feeling of learning (Carpenter 2017) may contribute to differences in learning. Another factor that might affect the learning outcomes in distributed practice research is the task that learners have to carry out during the sessions, as some research has found that more complex tasks may benefit more from massing than distribution (Suzuki and DeKeyser 2017). The next section brings together studies on an unexplored area of distributed practice effect that focuses on the task of viewing audiovisual materials.

Viewing time distribution

The number of studies that have incorporated distributed practice effects on L2 viewing is limited. These studies have taken one of two different angles. The first one is the perspective taken by previous reading research (e.g. Webb and Chang 2015), that looks at the items' distribution across the episodes.

The other angle is to approach the distribution practice effects by spacing or massing the viewings of the same or different episodes, given that authentic materials as TV series and movies cannot be easily modified in terms of target items' occurrence. In the first study, Ghebghoub (2021) looked at the items' distribution across the episodes and compared massed and spaced condition. The intermediate to upper-intermediate proficiency participants viewed eight one-hour episodes of a captioned documentary TV series over four sessions. The target vocabulary items were divided into two groups: massed in one session or spaced across the four sessions. Massed items occurred multiple times within one session only while spaced items occurred multiple times across all four of the sessions. The findings revealed no effect of item distribution on their learning; that is, whether the participants encountered the item repeated in the same session or throughout the four sessions did not affect the learning gains. This null effect of distribution of items occurrence transferred the findings of extensive reading-while-listening (Webb and Chang 2015) to captioned viewing.

On the other hand, Muñoz et al. (2022) looked at the effects of repeated viewing under massed and spaced viewing sessions, rather than massing or spacing target items across distributed viewing sessions. The B2 level English participants viewed an episode (21 minutes) of L2 captioned TV series twice in a row (massed condition) or with a week in between (spaced condition). The results indicated a tentative advantage of spacing over massing for learning the target vocabulary and multi-word expressions in the immediate post-test for meaning recognition, but not for meaning recall.

Finally, Avello and Muñoz (2022) explored the intensity of distribution of viewing sessions of different episodes. The study focused on vocabulary learning of primary school students (9 and 11 years old) from 11 episodes (11–22 minutes per episode) of an animated show. The participants watched L2 captioned episodes either twice a week or four times a week. The comparison between two age groups' written-word form recall revealed a significant effect of the viewing distribution for the younger group. The 9-year-old group had higher learning gains in the immediate post-test if they viewed the episodes four times a week rather than twice a week. As for the 11-year-old group, shorter or longer lags between the viewing sessions did not affect their scores. The authors suggested that the younger group was more sensitive to the effects of viewing distribution due to students' lower proficiency level, echoing Elgort and Warren's (2014) finding that lower proficiency level adults benefitted more from massed exposure. Similarly, Serfaty and Serrano (2022) found that for grammar learning, lower-proficiency learners profited more from shorter rather than longer spacing.

The study

The abovementioned studies made the first steps in the exploration of distributed practice effects in L2 viewing and yielded mixed results. While, in combination, the studies addressed both spaced and massed conditions, there was no experiment that explored the longer spaced, shorter spaced and massed viewing in a single intervention. Therefore, the aim of this study is to fill this gap and explore the effects of viewing time distribution of five full-length episodes of captioned TV series and see whether a massed condition (i.e. binge-watching) or longer or shorter spacing of episodes would lead to greater learning gains of MWUs.

In addition to exploring learning gains, this study explores two learner factors that might affect those gains: learners' proficiency level and feeling of learning. While learners' proficiency was identified as a significant factor in distributed practice effect in previous SLA research (e.g. Elgort and Warren 2014), the feeling of learning is a novel variable for time distribution research in SLA that might triangulate the learning outcome data by looking at the learners' perspective of massing and spacing the viewing sessions. In the field of cognitive psychology, it was found that in general learners believe massing is more beneficial than spacing, in contrast to the attested learning gains (see Carpenter 2017), which can consequently affect learning process from leisure viewing.

In light of all this, the following research questions are addressed in the study:

- (1) To what extent does viewing time distribution (i.e. longer spacing, shorter spacing and binge-watching) affect learning of MWUs from audiovisual input?
- (2) To what extent does proficiency level affect this learning?
- (3) What is the participants' perceived learning from the intervention and to what extent does it depend on the viewing time distribution?

Methodology

Participants

A total of 96 L1 Russian undergraduate engineering and information technology students participated in the study. Their proficiency varied from A2 to C1, with a mean of B2 according to the Common European Framework of Reference (CEFR, Council of Europe 2020). They were attending obligatory English classes 3 times a week (4.5 hours of instruction). Nine B1 + and B2 English university courses participated in the study. The intact classes had five different instructors that were trained beforehand by the first author to collect data. The classes were assigned to one of the three viewing conditions: longer spacing ($n = 26$), shorter spacing ($n = 40$) and massed binge-watching ($n = 30$). Care was made to guarantee that each viewing condition had classes from both B1 + and B2 courses.

Materials

Audiovisual input and target multiword units

The target TV series for this study was the American comedy show 'Fresh off the Boat' (Khan et al. 2015). This TV series has been used before in audiovisual input research (Majuddin et al. 2021; Muñoz et al. 2022; Pujadas and Muñoz 2019). The participants were exposed to five episodes of the first season of the TV series (109 minutes of audiovisual input). None of the participants reported having viewed the TV show before. Based on the TV series episodes' scripts, 25 target MWUs were identified in consultation with the course teachers based on the MWUs' likelihood to be unfamiliar to the participants (B1 + and B2 English courses). The MWUs that mostly appeared at least twice throughout the five episodes were selected (see Table 1). The MWUs in this study are defined as expressions that are longer than a single word (Majuddin et al. 2021), and fully filled (fixed) language constructions that have no variability in the target input (the TV series episodes) (Pattemore and Muñoz 2022).

Vocabulary Knowledge Scale pre-/post-test

The target multiword units were tested by an adaptation (e.g. Bardovi-Harlig 2014) of the standardised Vocabulary Knowledge Scale (VKS) (Wesche and Paribakht 1996). The original VKS tested the depth of vocabulary knowledge, while the adaptation targets the knowledge of MWUs. The

Table 1. Target multi-word units.

| MWU | Total frequency of occurrence | MWU | Total frequency of occurrence |
|-------------------------|-------------------------------|-------------------------------|-------------------------------|
| To figure out | 7 | Change the game | 2 |
| To be a good cop | 4 | Drop the mic | 2 |
| To take time | 4 | Find someone to identify with | 2 |
| A hate crime | 3 | To be on one's side | 2 |
| Bring it! | 3 | To get a nut | 2 |
| Do whatever it takes | 3 | To plug something in | 2 |
| Fresh off the boat | 3 | To put down an animal | 2 |
| Next thing you know | 3 | Show somebody around | 2 |
| Look out for each other | 3 | Stand up for someone | 2 |
| To dine and dash | 3 | Wield the hammer | 2 |
| To fit in | 3 | A slippery slope | 1 |
| To run out on the bill | 3 | To do your best | 1 |
| To project an image | 3 | | |

VKS consists of a self-reporting elicitation scale (see [Figure 1](#)) where learners can identify the level of MWU's familiarity and knowledge, from initial recognition to ability to use it in a context. The test is designed to get students to indicate their degree of knowledge of the target items. It is an instrument of initial recognition and use of new MWUs. It elicits both self-perceived and demonstrated knowledge (Wesche and Paribakht 1996). The VKS in this study included the term 'expression' to facilitate participants' comprehension as they may not have been familiar with the term 'multi-word unit'.

The pre-test consisted of 50 items, 25 target items and 25 distractors to divert students' attention from the target MWUs. The post-test included only the target items. The pre-test was piloted beforehand with a comparable group of participants ($n = 10$) from the same university to assure that the target items were appropriate for the B1 + and B2 language courses. To score the results, the guidelines provided by Wesche and Paribakht (1996) were used. If a student chose option '1', that indicated that they were not familiar with the expression at all, they received 1 point. If a student chose option 2 indicating that the expression was familiar, but its meaning was unknown, the participant received 2 points. If the participant chose options 3 or 4, and provided a correct synonym, translation or definition, they received 3 or 4 points respectively. If a student provided an incorrect synonym, translation or definition, they received 2 points. The participants received the maximum 5 points if they used the MWU appropriately in a sentence and provided its synonym, translation or definition as well. If the participants did not provide an appropriate example in a context, but provided a correct synonym, translation or definition, they received 4 points. If the synonym, translation or definition was not correct either, the student received 2 points. Following this scoring procedure, the minimum score for 25 items was 25 points and the maximum score was 125 points. Two inter-raters scored the tests until agreement was reached.

Proficiency test

Participants' proficiency level was measured by the Oxford Placement Test (Allan 2004), a standardised proficiency test that consists of two parts: listening and grammar. The total score (maximum 200) for the two parts can be converted into a CEFR grid. Participants were assigned to proficiency level groups following the OPT guidelines: A2 elementary level from 104 to 119 points, B1 low-intermediate level from 120 to 134, B2 upper-intermediate level from 135 to 149 and C1 advanced level from 150 to 171.

Post-viewing activities

To direct participants' attention to the content of the video and measure their overall comprehension, they were asked to answer five comprehension questions after each episode, a mixture of multiple-choice and open-ended. At the end of the last, fifth episode, the participants were also asked to answer an open-ended question about their feeling of learning from the five episodes. The

*Please provide your answer (1, 2, 3, 4, 5) for the expression **a slippery slope***

1. I have never seen this expression before (I don't remember).
2. I have seen this expression before, but I don't know what it means.
3. I have seen this expression before, and I think it means:
_____ (synonym, translation, or definition)
4. I know this expression, it means:
_____ (synonym, translation, or definition)
5. I can use this expression in a context:
_____ (write a sentence)

(If you do Section 5, please also do Section 4)

Figure 1. Vocabulary Knowledge Scale adapted from Wesche and Paribakht (1996).

Table 2. VKS scores.

| | Pre-test | | | Post-test | | |
|-------------------------------|---------------|-----|-----|---------------|-----|-----|
| | Mean (SD) | Min | Max | Mean (SD) | Min | Max |
| All participants ($n = 96$) | 60.32 (12.71) | 38 | 105 | 72.79 (15.61) | 46 | 115 |
| Longer spacing ($n = 26$) | 59.23 (12.40) | 38 | 84 | 78.11 (15.99) | 51 | 100 |
| Shorter spacing ($n = 40$) | 61.00 (11.17) | 44 | 88 | 71.10 (14.23) | 47 | 102 |
| Binge-watching ($n = 30$) | 60.36 (15.07) | 39 | 105 | 70.43 (16.44) | 46 | 115 |

participants were given the following prompt: ‘Do you feel you have learnt something from this TV series? (e.g. vocabulary, expressions, grammar, pronunciation, sociocultural references, or nothing)’. The students listed as many categories as they felt they had learnt.

Procedure

All participants filled out the consent form, completed the proficiency test (60 minutes) and the pre-test (about 40 minutes) five days before viewing the first episode. The longer spacing group watched one episode per week over a period of five weeks (always on the same day of the week, inter-session interval (ISI) of 7 days), the shorter spacing group viewed all five episodes in one week (one episode per day from Monday to Friday, ISI of 1 day) and the massed binge-watching group watched all five episodes in one session. The posttest took place immediately after viewing the last episode.

The participants completed the proficiency test and the pre-/post-tests in the classroom. Spacing groups’ participants watched episodes one to four at home, while episode five was watched in the classroom because it was followed by the posttest. The binge-watching group watched all the episodes in the classroom as they completed the post-test immediately after viewing the five episodes in one session. Students received individual links to the episodes through the EdPuzzle platform (Sabrià 2013). This web platform allows researchers and teachers to monitor the viewing progress in real time and check viewing completion. It also blocks the possibility of rewinding or fast-forwarding the video and stops playing the video if the viewer opens another tab on their browser. In addition, the platform can include questions at the end of the video. The participants answered the comprehension questions and the feeling of learning question through this platform. The intervention was implemented as a part of the course and the participants received class credits for viewing the episodes and completing the tests. They were not aware of the nature of the experiment and were debriefed at the end of the intervention.

Results

Preliminary analysis

The VKS scores’ descriptive statistics per group are presented in Table 2. The test scores were not normally distributed, therefore the log transformation was applied. Log transformed scores were normally distributed as the values of Shapiro–Wilk and Kolmogorov–Smirnov tests were non-significant. The preliminary analysis showed that there was no difference in the pretest scores ($F(2,93)$

Table 3. Oxford placement test scores.

| | Mean (SD) | Min | Max |
|-------------------------------|----------------|-----|-----|
| All participants ($n = 96$) | 139.71 (16.81) | 104 | 171 |
| Longer spacing ($n = 26$) | 134.47 (14.51) | 108 | 166 |
| Shorter spacing ($n = 40$) | 145.40 (15.01) | 109 | 171 |
| Binge-watching ($n = 30$) | 139.55 (18.81) | 104 | 169 |

Table 4. Pre-/post-test scores' per proficiency group.

| | | Mean (SD) | Min | Max |
|-------------|-----------|---------------|-----|-----|
| A2 (n = 10) | Pre-test | 52.60 (10.86) | 41 | 73 |
| | Post-test | 62.20 (6.14) | 53 | 73 |
| B1 (n = 34) | Pre-test | 55.17 (9.95) | 38 | 78 |
| | Post-test | 67.17 (14.33) | 46 | 100 |
| B2 (n = 26) | Pre-test | 62.50 (9.43) | 50 | 84 |
| | Post-test | 75.80 (15.39) | 51 | 100 |
| C1 (n = 26) | Pre-test | 67.84 (15.10) | 39 | 105 |
| | Post-test | 81.19 (15.28) | 56 | 115 |

= .150, $p = .861$, $\eta p^2 = .003$) between the three groups. As for the proficiency levels, students were assigned to a proficiency group following the OPT guidelines: A2, B1, B2 and C1. The analysis showed that the three groups' overall proficiency levels were comparable ($F(2,93) = .807$, $p = .449$, $\eta p^2 = .017$). Table 3 shows the mean proficiency scores per each viewing condition Table 4 shows the pre-/ and post-test scores per each proficiency level.

Main analysis

To answer the first and the second research questions, a series of Generalized Linear Mixed Models (GLMMs) were fitted in SPSS software (version 27). The effect sizes for Cohen's d were interpreted using the following criteria (Plonsky and Oswald 2014): small ($d = .40$), medium ($d = .70$) and large ($d = 1.00$). The effect sizes for partial eta squared (ηp^2) were considered as small ($\eta p^2 = .01$), medium ($\eta p^2 = .06$) and large ($\eta p^2 = .014$) following Cohen's (1988) guidelines.

The first model included the scores as a dependent variable, and time (pretest or posttest), group (longer spacing, shorter spacing, binge-watching), and proficiency level (A2, B1, B2, C1) as factors. The model also included three interactions between group and time, group and proficiency, and group, time and proficiency. Time was included as a repeated measure.

The model yielded no significant effect of the interaction between group and proficiency ($F(6, 168) = 1.691$, $p = .126$, $\eta p^2 = .056$), or group, time and proficiency ($F(9, 168) = .263$, $p = .984$, $\eta p^2 = .013$), suggesting the difference between the three viewing conditions scores' did not depend on the proficiency level. These interactions were removed and the model was fitted again.

The final model showed a significant effect of time ($F(1, 183) = 60.666$, $p < .001$, $\eta p^2 = .248$), group ($F(2, 183) = 4.441$, $p = .013$, $\eta p^2 = .046$), proficiency level ($F(3, 183) = 16.133$, $p < .001$, $\eta p^2 = .209$) and a significant interaction between time and group ($F(2, 183) = 4.031$, $p = .019$, $\eta p^2 = .042$).

The pairwise comparison with sequential Bonferroni correction revealed that the participants scored significantly higher in the post-test than in the pre-test ($t(183) = 7.789$, $p < .001$, $d = 1.15$). Regarding the group and time interaction, the pairwise comparison showed that there was a significant effect of group at post-test (see Figure 2). The participants in the longer spacing group scored significantly higher than the shorter spacing group ($t(183) = 3.652$, $p = .001$, $d = .54$) and the binge-watching group ($t(183) = 3.694$, $p = .001$, $d = .55$). There was no significant difference between the shorter spacing and binge-watching groups' scores ($t(183) = .301$, $p = .764$, $d = .04$).

Finally, the last significant factor in the model was proficiency level (see Figure 3). The pairwise comparison indicated that the participants with higher proficiency level outperformed those with the lower language proficiency, except for the non-significant difference between the A2 and B1 proficiency groups ($t(183) = -1.022$, $p = .308$, $d = .15$) and the B2 and C1 proficiency groups ($t(183) = -2.093$, $p = .075$, $d = .31$). The rest of the comparisons between the lower and higher proficiency levels were statistically significant (A2 and B2 ($t(183) = -3.555$, $p = .001$, $d = .52$); A2 and C1 ($t(183) = -5.179$, $p < .001$, $d = .76$); B1 and B2 ($t(183) = -3.723$, $p = .001$, $d = .55$); B1 and C1 ($t(183) = -5.909$, $p < .001$, $d = .87$)).

The third research question was concerned with the participants' perceptions of learning from the interventions. The answers from the open-ended questionnaire were coded into six categories

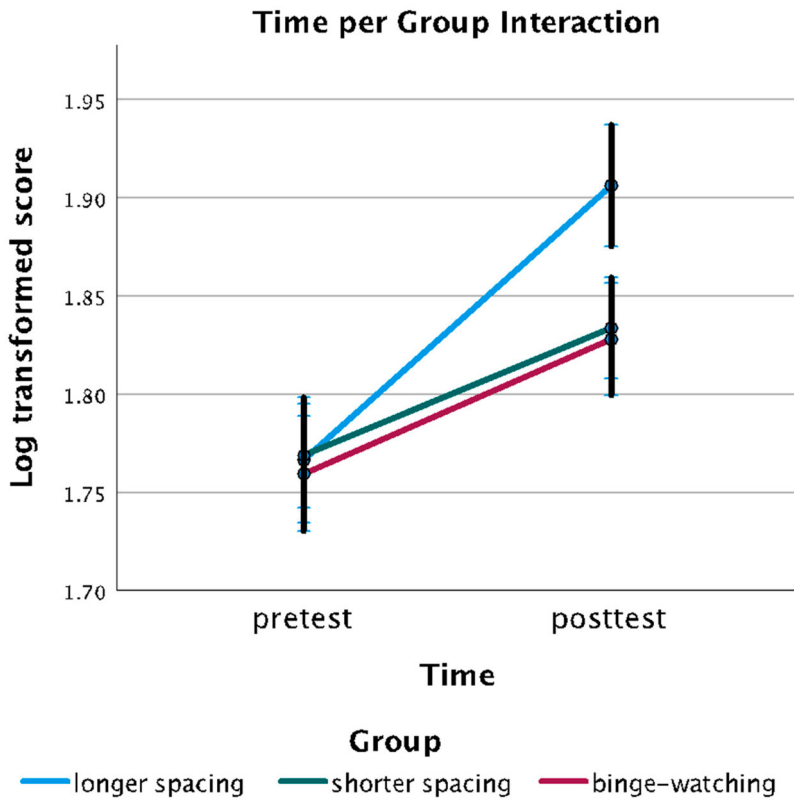


Figure 2. Interaction between time and group.

(vocabulary, expressions, grammar, pronunciation, sociocultural aspects and no feeling of learning). The descriptive exploration of data revealed that 93% of the participants reported benefitting from the intervention through language learning, while only 7% of the participants said that they did not learn anything from the TV series. Regarding which language features were perceived as learnt, the participants could report as many language features as possible. The results showed that 39% of participants felt they learnt pronunciation, 37% vocabulary, 29% expressions (MWUs), 21% sociocultural aspects from the TV series and 10% grammar, see Table 5 for the responses per viewing group.

To see whether this feeling of learning depended on the participants' viewing time distribution, a Binomial Linear Model for Repeated Measures was fitted with a binary response (yes or no) as a dependent variable, and feeling of learning category (vocabulary, expressions, grammar, pronunciation, sociocultural aspects and no feeling of learning) and viewing time distribution (longer spacing, shorter spacing and binge-watching) as factors. The analysis revealed no significant interaction between the group and feeling of learning category ($F(12, 469) = .905, p = .542, \eta p^2 = .022$), suggesting that different viewing time distributions did not affect participants' feeling of learning.

Discussion

Learning from audiovisual input and viewing distribution

The present study was carried out with the aim of exploring the uptake of MWUs from three types of viewing distribution: longer spacing (7 day ISI), shorter spacing (1 day ISI) and massed binge-watching. The results of the first research question indicated that the participants significantly improved their knowledge of the target items between the pre- and post-test regardless of their viewing

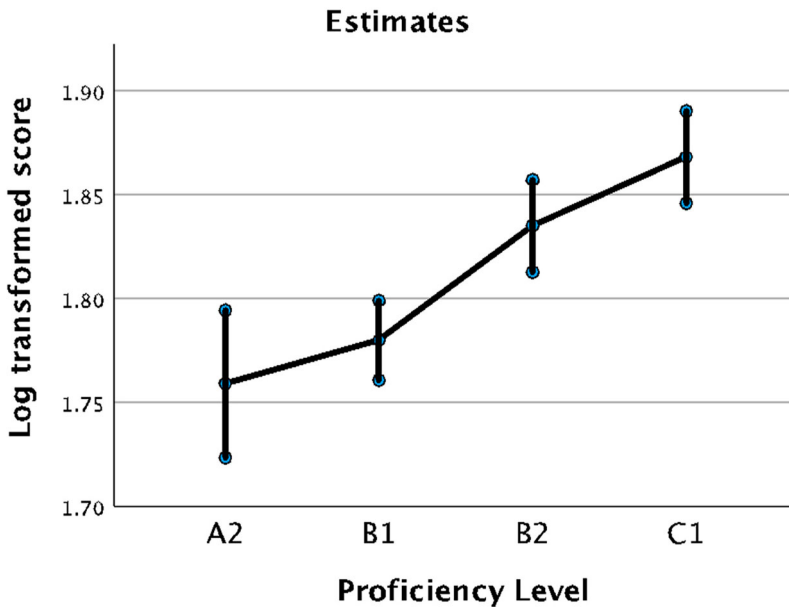


Figure 3. Effect of proficiency on scores.

distribution. This goes in line with the previous research on MWU uptake from captioned audiovisual input (Majuddin et al. 2021) and provides additional evidence that captioned audiovisual input is beneficial not only for vocabulary (see Montero Perez 2022), but also for MWU learning.

Regarding the viewing time distribution, the results yielded that longer spacing had an advantage over both shorter spacing and massed binge-watching. The advantage of the longer spacing group over the massed group goes in line with the previous psychology and SLA research where spacing was more beneficial for learning than massing (Carpenter 2017; Kim and Webb 2022). This result also goes well with Muñoz et al. (2022) who found a tendency for the spaced group to outperform the massed group after repeated viewing. This result suggests that retrieval effort in longer spacing might help retention by providing separate learning episodes and encoding variability compared to the massed condition. It is also possible to assume that longer spacing, compared to massed spacing, enhances attention and therefore promotes learning.

Considering the advantage of longer spacing over shorter spacing in this study, it confirms the lag effect (Cepeda et al. 2008), but clashes with the SLA results where shorter spacing was more beneficial in the immediate posttest, but longer spacing had a greater advantage for long-term retention demonstrated by a delayed posttest (e.g. Serrano and Huang, 2018). Our results go in contrary with this as the shorter spacing group did not outperform the longer spacing group in the posttest that took place immediately after the last episode viewing. This difference could be explained by the differences between the previous studies on distributed practice effect and the present study. For example, Serrano and Huang (2018) looked at repeated reading-while-listening (bi-modal input)

Table 5. Feeling of learning from the intervention.

| | Longer spacing | Shorter spacing | Binge-watching |
|------------------------|----------------|-----------------|----------------|
| Vocabulary | 60% | 31% | 36% |
| Expressions (MWUs) | 10% | 23% | 44% |
| Grammar | 0% | 11% | 12% |
| Pronunciation | 30% | 37% | 44% |
| Sociocultural aspects | 0% | 26% | 24% |
| No feeling of learning | 20% | 9% | 0% |

of the same text, while in our intervention participants were exposed to different TV episodes (multi-modal input) during each session. It is possible that when considering repeated exposures in different contexts (episodes), rather than repeated exposure to the same materials (Serrano and Huang, 2018), the immediate effect of shorter spacing does not appear. More research on viewing comparisons with repeated or non-repeated audiovisual stimuli, could shed light on this matter. Regarding the type of knowledge that was tested in our study, the findings are in line with Elgort et al. (2018) who suggested that better learning of meaning occurs in the spaced manner. The results are also coincidental with studies looking at the lag effect on grammar, in which long spacing is more effective for explicit knowledge than short spacing (Bird 2010). Our test could be claimed to elicit explicit rather than implicit knowledge, and therefore we could tentatively suggest that longer spacing is beneficial for explicit learning of MWUs as well.

Surprisingly, the shorter spacing group and the massed binge-watching group performed similarly in this study. Although it goes in contrary with the majority of the previous research suggesting an advantage of short spacing over massing (Serrano 2022), it is possible that one day spacing between the viewing sessions was too short for memory consolidation to appear. In addition, the retrieval effort that is associated with spacing conditions, comes at work when the task is challenging enough and the information about the target items is not too recent. Again, a one-day inter-session interval might not have been sufficient to implement the necessary cognitive processes to benefit from spacing to a greater extent. From the audiovisual input perspective, a plausible explanation for the lack of difference between the massed binge-watching and shorter spacing groups may be attributed to the exposure to captioned rather than uncaptioned TV series. Previous research suggests that captions have a mitigating role in learning from audiovisual input by levelling the cognitive individual differences (Pattemore and Muñoz 2020). Although binge-watching might be a more challenging task as it could lead to deficient processing and lack of attention (Carpenter 2017; Steins-Loebers et al. 2020), it is possible that captions attenuated the limits of massed exposure and eased the processing load of the binge-watching group to the extent that they performed at the same level as the shorter spacing group. An interesting trajectory for future research could be exploring the effects of viewing time distribution on uncaptioned input processing as well. As both binge-watching and uncaptioned input imply extra difficulty, it is possible that the advantage of shorter spacing over massing would appear.

Viewer factors

The second research question analysed the possible effect of viewers' proficiency on learning MWUs from three types of viewing time distribution. The results indicated that there was a significant effect of proficiency on learning, confirming previous research suggesting that those who already had higher levels of proficiency would benefit more than those with lower levels (e.g. Suárez and Gesa 2019). Interestingly, there was no significant interaction between proficiency level and viewing time distribution, suggesting that successful learning from one type of distribution or another does not depend on learners' proficiency, at least not with this type of material. This finding does not support Avello and Muñoz's (2022) finding that lower proficiency learners benefited from shorter lags more. It is important to keep in mind the age differences in the two experiments, as age was another significant factor in the distributed practice research of Kim and Webb (2022). That study explored the effects of viewing time distribution on children while the present study targeted adult learners. Therefore, it is possible that age is a stronger factor than proficiency in viewing time distribution. In addition, the participants in the present study might have reached the necessary proficiency level (intermediate) to benefit from captioned input to the full extent (Danan 2004). Alternatively, the availability of captions might have made the learning task more feasible and supported the lower proficiency viewers in the massed condition as well.

The last research question looked at the participants' views of learning from the intervention. Given the possibility of deficient processing and attention decrease during exposure to massed

input, it would have been logical to expect lower levels of feeling of learning in the binge-watching condition. Although the results of the model indicated that the three groups did not differ in their perceived learning from the TV series, the descriptive statistics (see [Table 5](#)) suggest a tendency for the binge-watching group to perceive more learning from the intervention, especially when comparing with the longer spacing group. This goes in line with Birnmbaum et al. (2013) where participants felt they learnt more from massed than spaced condition. These results confirm that learners on their own might not be able to detect which viewing distribution is more beneficial for their learning. Nevertheless, it is premature to make conclusions based on this result as the participants were each exposed to only one type of viewing distribution. Therefore, future research could look at the feeling of learning in a within-subjects design when the same participants are exposed to various viewing distributions.

Limitations

This study is not without limitations. One of them lies in the absence of a delayed posttest, caused by participant drop out. On the other hand, although delayed posttest results may have provided insights about the retention of the MWUs, delayed posttest inclusion could also lead to additional exposure to the target items and increase the possibility of learning outside of the intervention by providing additional practice and raising attention to the target MWUs. The second limitation lies in the absence of a control group that would not watch the episodes and only complete the tests to provide greater evidence that learning took place due to the viewing. The inclusion of a control group implies a practical difficulty for the present study's design as it involves three control and three experimental groups (longer spacing, shorter spacing and massed). It is a challenging task in classroom-based research where access to a large number of comparable participants might be limited.

Conclusion

This study expands time distribution research into the area of learning from audiovisual input and provides an original contribution by comparing learning gains from different viewing distributions of captioned TV series. Previous time distribution research mainly focused on distribution of items occurrence rather than learning sessions distribution. While research on distribution of items occurrence is valid for creation of educational videos, a logical and ecologically valid trajectory for the OV audiovisual input research is to approach the distribution practice effects by spacing or massing the viewings of the same or different episodes. Therefore, this is the first study to compare longer spacing, shorter spacing and massing within the same intervention. Although the results indicated that learning of MWUs could take place regardless of the viewing distribution group, it appears advisable for L2 learners and instructors to apply longer spacing of the viewing sessions to facilitate learning and achieve better outcomes. Especially if we want to promote efficient, autonomous learning beyond the classroom.

Note

1. Although we refer to binge-watching as a massed condition, it does not strictly follow the traditional view of massing items in cognitive psychology (e.g. Carpenter 2017). Massed conditions typically involve exact subsequent repetitions of stimuli, while binge-watching implies sequential viewing of episodes but where the target items may be spaced in their distribution within the episodes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by grant PID2019-110594GB-I00 from the Spanish Ministry of Science and Innovation, and EuroSLA 2021 early career researcher award.

ORCID

Anastasia Pattemore  <http://orcid.org/0000-0003-2038-8017>

Carmen Muñoz  <http://orcid.org/0000-0002-7001-4155>

References

- Allan, D. 2004. *Oxford Placement Test*. Oxford: Oxford University Press.
- Avello, D., and C. Muñoz. 2022. Vocabulary learning through captioned-video viewing: the role of treatment- and word-related factors. Paper presented at the BAAL Vocabulary Special Interest Group's Conference, Exeter, July 18-19.
- Bardovi-Harlig, K. 2014. Awareness of meaning of conventional expressions in second-language pragmatics. *Language Awareness* 23, no. 1-2: 41-56. doi:10.1080/09658416.2013.863894.
- Bird, S. 2010. Effects of distributed practice on the acquisition of second language English syntax. *Applied Psycholinguistics* 31, no. 4: 635-650. doi:10.1017/S0142716410000172.
- Birnbaum, M.S., N. Kornell, E.L. Bjork, and A.R. Bjork. 2013. Why interleaving enhances inductive learning: the roles of discrimination and retrieval. *Memory & Cognition* 41: 392-402. doi:10.3758/s13421-012-0272-7.
- Carpenter, S.K. 2017. Spacing effects on learning and memory. In *In Learning and Memory: A Comprehensive Reference*, ed. J.H. Byrne, 465-485. Elsevier. doi:10.1016/B978-0-12-809324-5.21054-7
- Castro, D., J.M. Rigby, D. Cabral, and V. Nisi. 2021. The binge-watcher's journey: investigating motivations, contexts, and affective states surrounding Netflix viewing. *Convergence* 27, no. 1: 3-20. doi:10.1177/1354856519890856.
- Cepeda, N.J., E. Vul, D. Rohrer, J. T. Wixted, and H. Pashler. 2008. Spacing effects in learning: a temporal ridgeline of optimal retention. *Psychological Science* 19, no. 11: 1095-1102. doi:10.1111/j.1467-9280.2008.02209.x.
- Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. New York: Lawrence Erlbaum Associates.
- Council of Europe. 2020. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume*. Strasbourg: Council of Europe Publishing.
- Danan, M. 2004. Captioning and subtitling: undervalued language learning strategies. *Meta* 49, no. 1: 67-77. doi:10.7202/009021ar.
- De Wilde, V., M. Brysbaert, and J. Eyckmans. 2021. Formal versus informal L2 learning: how do individual differences and word-related variables influence French and English L2 vocabulary learning in Dutch-speaking children? *Studies in Second Language Acquisition*, 1-25. doi:10.1017/S0272263121000097.
- Elgort, I., M. Brysbaert, M. Stevens, and E. Van Assche. 2018. Contextual word learning during reading in a second language: an eye-movement study. *Studies in Second Language Acquisition* 40, no. 2: 341-366. doi:10.1017/S0272263117000109.
- Elgort, I., and P. Warren. 2014. L2 vocabulary learning from reading: explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning* 64, no. 2: 365-414. doi:10.1111/lang.12052.
- Frumuselu, A.D., S. De Maeyer, V. Donche, and M.C. Plana. 2015. Television series inside the EFL classroom: bridging the gap between teaching and learning informal language through subtitles. *Linguistics and Education* 32: 107-117. doi:10.1016/j.linged.2015.10.001.
- Gheghoub, S. 2021. Imagery in L2 captioned video: investigating incidental vocabulary learning from extensive viewing as a function of modality, contiguity, and spacing. PhD diss., University of York. <https://etheses.whiterose.ac.uk/30597/>.
- Khan, N., J. Kasdar, M. Melvin, R. Blomquist, E. Huang, and J. McEwen. 2015. *Fresh off the Boat TV Series*. Los Angeles, CA: ABC.
- Kim, S. K., and S. Webb. 2022. The effects of spaced practice on second language learning: a meta-analysis. *Language Learning* 72, no. 1: 269-319. doi:10.1111/lang.12479.
- Küpper-Tetzel, C. E. 2014. Understanding the distributed practice effect: strong effects on weak theoretical grounds. *Zeitschrift für Psychologie* 222, no. 2: 71-81. doi:10.1027/2151-2604/a000168.
- Lindgren, E., and C. Muñoz. 2013. The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism* 10, no. 1: 105-129. doi:10.1080/14790718.2012.679275.
- Majuddin, E., A. Siyanova-Chanturia, and F. Boers. 2021. Incidental acquisition of multiword expressions through audio-visual materials. *Studies in Second Language Acquisition* 43, no. 5: 985-1008. doi:10.1017/S0272263121000036.
- Martin, E. 1968. Stimulus meaningfulness and paired-associate transfer: an encoding variability hypothesis. *Psychological Review* 75, no. 5: 421-441. doi:10.1037/h0026301.

- Mayer, R. E., L. Fiorella, and A. Stull. 2020. Five ways to increase the effectiveness of instructional video. *Educational Technology Research and Development* 68, no. 3: 837–852. doi:10.1007/s11423-020-09749-6.
- Montero Perez, M. 2022. Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching* 55, no. 2: 163–192. doi:10.1017/S0261444821000501.
- Muñoz, C. 2020. Boys like games and girls like movies. Age and gender differences in out-of-school contact with English. *RESLA, Spanish Journal of Applied Linguistics* 33, no. 1: 171–201. doi:10.1075/resla.18042.mun.
- Muñoz, C., A. Pattenmore, and D. Avello. 2022. Exploring repeated captioning viewing as a way to promote vocabulary learning: time lag between repetitions and learner factors. *Computer Assisted Language Learning*.
- Pattenmore, A., and C. Muñoz. 2020. Learning L2 constructions from captioned audio-visual exposure: the effect of learner-related factors. *System* 93. doi:10.1016/j.system.2020.102303.
- Pattenmore, A., and C. Muñoz. 2022. Captions and learnability factors in learning grammar from audio-visual input. *The JALT CALL Journal* 18, no. 1: 83–109. doi:10.29140/jaltcall.v18n1.564.
- Plonsky, L., and F. L. Oswald. 2014. How big is 'big'? Interpreting effect sizes in L2 research. *Language Learning* 64, no. 4: 878–912. doi:10.1111/lang.12079.
- Puimège, E., and E. Peters. 2019. Learning L2 vocabulary from audiovisual input: an exploratory study into incidental learning of single words and formulaic sequences. *The Language Learning Journal* 47, no. 4: 424–438. doi:10.1080/09571736.2019.1638630.
- Pujadas, G., and C. Muñoz. 2019. Extensive viewing of captioned and subtitled TV series: a study of L2 vocabulary learning by adolescents. *The Language Learning Journal* 47, no. 4: 479–496. doi:10.1080/09571736.2019.1616806.
- Rubeking, B., and C.C. Bracken. 2021. Binge watching and serial viewing: comparing new media viewing habits in 2015 and 2020. *Addictive Behaviors Reports* 14: 100356. doi:10.1016/j.abrep.2021.100356.
- Sabrià, Q. 2013. Edpuzzle. <https://edpuzzle.com/>.
- Serfaty, J., and R. Serrano. 2022. Lag effects in grammar learning: a desirable difficulties perspective. *Applied Psycholinguistics* 43, no. 3: 513–550. doi:10.1017/S0142716421000631.
- Serrano, R. 2012. Is intensive learning effective? Reflecting on the results from cognitive psychology and the second language acquisition literature. In *Intensive Exposure in Second Language Learning*, ed. C. Muñoz, 3–22. Bristol: Multilingual Matters. doi:10.21832/9781847698063-004
- Serrano, R. 2022. A state-of-the-art review of distribution-of-practice effects on L2 learning. *Studies in Second Language Learning and Teaching* 12, no. 3: 355–379. doi:10.14746/ssllt.2022.12.3.2.
- Serrano, R., and H.Y. Huang. 2018. Learning vocabulary through assisted repeated reading: how much time should there be between repetitions of the same text? *TESOL Quarterly* 52, no. 4: 971–994. doi:10.1002/tesq.445.
- Serrano, R., and H. Huang. 2021. Time distribution and intentional vocabulary learning through repeated reading: a partial replication and extension. *Language Awareness*. doi:10.1080/09658416.2021.1894162.
- Steins-Loeber, S., T. Reiter, H. Averbek, L. Harbarth, and M. Brand. 2020. Binge-watching behaviour: the role of impulsivity and depressive symptoms. *European Addiction Research* 26: 141–150. doi:10.1159/000506307.
- Suárez, M. M., and F. Gesa. 2019. Learning vocabulary with the support of sustained exposure to captioned video: do proficiency and aptitude make a difference? *The Language Learning Journal* 47, no. 4: 497–517. doi:10.1080/09571736.2019.1617768.
- Suzuki, Y., and R. DeKeyser. 2017. Effects of distributed practice on the proceduralization of morphology. *Language Teaching Research* 21, no. 2: 166–188. doi:10.1177/1362168815617334.
- Toppino, T. C., and E. Gerbier. 2014. About practice: repetition, spacing, and abstraction. In *The Psychology of Learning and Motivation*, ed. B. H. Ross, 113–189. Elsevier Academic Press. doi:10.1016/B978-0-12-800090-8.00004-4
- Vanderplank, R. 2016. *Captioned Media in Foreign Language Learning and Teaching: Subtitles for the Deaf and Hard-of-Hearing as Tools for Language Learning*. Oxford: Palgrave Macmillan.
- Webb, S. 2014. Extensive viewing: language learning through watching television. In *Language Learning Beyond the Classroom*, eds. D. Nunan and J.C. Richards, 159–168. New York: Routledge.
- Webb, S., and A. C-S. Chang. 2015. Second language vocabulary learning through extensive reading with audio support: how do frequency and distribution of occurrence affect learning? *Language Teaching Research* 19, no. 6: 667–686. doi:10.1177/1362168814559800.
- Wesche, M., and T.S. Paribakht. 1996. Assessing second language vocabulary knowledge: depth vs breadth. *The Canadian Modern Language Review* 53, no. 1: 13–40. doi:10.3138/cmlr.53.1.13.