I Introduction

One of the potential advantages of Study Abroad (SA) as a context for second language (L2) learning is the provision of large amounts of authentic input. Research has shown multiple benefits of SA (e.g., Llanes, 2011; Sanz and Morales-Front, 2018) but it has also shown wide differences across studies as well as across sojourners sharing apparently similar contextual conditions in the same study. Together with during-SA L2 contact, pre-SA proficiency level has been examined in an attempt to understand such inter-individual variability in linguistic development during SA (e.g., Faretta-Stutenberg and Morgan Short, 2018). Research results are mixed, with some studies showing greater gains the lower the learners' initial L2 level (e.g. Brecht et al., 1990), and others showing greater gains with higher pre-departure language skills (e.g., Cubillos et al., 2008). These differences may be partly explained by variable measurement choice (Tullock and Ortega, 2017), and because for some aspects there may just be more room for development. DeKeyser (2007: 211) notes the explanations offered for why the most advanced students may learn more: they interact more (Brecht and Robinson, 1993); they are more adept at managing the ceaseless flow of L2 input (Rivers, 1998: 492); and they take more frequent advantage of extracurricular listening (e.g., to radio, films, television) (Segalowitz and Freed, 2004: 195). It follows that learners' preparation at home should include abundant exposure to authentic spontaneous L2 speech (DeKeyser, 2010; Paige et al., 2002), which is often lacking in the traditional classroom curriculum.

This paper will examine the extent to which viewing audio-visual material in the target language can support learners' vocabulary and grammar development in preparation for a sojourn abroad and during SA. The notion that the combination of images and verbal information in audio-visual input may enhance L2 learning (Sydorenko, 2010) is supported by information processing theories sustaining that the presentation of words with pictures allows learners to make connections between them (Mayer, 2014). The benefits of such multimodal input may be increased by the addition

of caption text with which the information is distributed among three channels, easing the load of working memory (Vanderplank, 2016: 147). Moreover, audio-visual material complies with Nation's (2007) five conditions for suitable input: it is processed in large quantities, it is comprehensible, it is engaging (Webb, 2010), it is familiar to the language learners, and it provides contextual cues (i.e. through image and dialogue) (Rodgers, 2013). Audio-visual input can positively help learners to get used to fast authentic speech inside and outside the classroom, train their listening skills and practice effective listening strategies (Vandergrift, 2007). It can also more generally contribute to improve learners' proficiency level, as indicated by mounting research evidence in the areas of vocabulary (e.g., Pujadas and Muñoz, 2019), comprehension (e.g., Montero Perez et al., 2013), pronunciation (Wisniewska and Mora, 2020), and grammar (e.g., Lee and Révész, 2018; Pattemore and Muñoz, 2020).

A distinctive advantage of audio-visual viewing is that it can occur anywhere and anytime. It can begin in the classroom and continue outside its walls (Collins and Muñoz, 2016). In fact, one of the principles of *extensive* viewing (Webb and Rodgers, 2009a), as a fitting and complementary alternative to extensive reading (Nation, 2015) in the digital area, is that classroom-based extensive viewing should guide out-of-school viewing (Webb, 2015). Moreover, extensive viewing provides unlimited opportunities for incidental learning in a learner-regulated manner. In particular, studies have shown the potential of audio-visual input for incidental vocabulary learning due to its lexical richness and repeated encounters with low frequency words (Rodgers, 2018; Sydorenko, 2010; Webb and Rodgers, 2009b). Similarly, Rodgers and Webb (2011) recommend *narrow* viewing based on the benefits of narrow reading (Schmitt and Carter, 2000) for comprehension and learning. When viewing episodes from the same TV series, learners accumulate background knowledge which helps their top-down processing and their familiarization with the language used by the characters (accent, vocabulary and grammar).

This paper focuses on vocabulary and grammar learning through viewing audio-visual input and, in particular, on the impact of two characteristics of TV series that have the potential of helping language learners. First, words tend to reoccur in the same season of a TV programme (Rodgers and Webb, 2011), and grammatical constructions may reoccur with higher frequency than in unrelated programmes as well. Second, audio-visual material may display on-screen text, subtitles in the L1 or captions in the L2, that help content comprehension and have the potential to enhance language

learning as learners are exposed to a large amount of input simultaneously through image, text, and sound.

II Background

1 Frequency effects in learning vocabulary and grammatical constructions

Frequency is among the most important predictors of vocabulary learning, even if not the only one (see Hashimoto and Egbert, 2019). Frequency effects in audio-visual input have been especially examined in vocabulary studies and it has been found that repeated encounters with unknown words have a positive effect on learning those words (e.g., Peters et al., 2016; Peters, 2019; Peters and Webb, 2018; Rodgers, 2013; Rodgers and Webb, 2019). However, the precise number of encounters needed to learn a word incidentally is difficult to establish, and it may depend to some extent on the input mode (written or aural). For instance, reading research has suggested between 8 and 10 encounters, although not all aspects of word knowledge are learned in the same way (e.g., Chen and Truscott, 2010; Pellicer-Sánchez, 2016). Listening research has suggested that frequency of encounters has a smaller effect in listening than in reading (Van Zeeland and Schmitt, 2013; Vidal, 2011) and that listening requires more encounters with words to learn them given the transient nature of the mode; although if listening is supported by gesture, as few as 8 occurrences may be sufficient for learning after minimal exposure (Gullberg et al., 2010). In reading-while-listening, Brown et al., (2008) suggested that between 7 and 9 encounters were needed. In viewing research, it has been suggested that fewer encounters may be required because of the presence of visual support (Webb and Rodgers, 2009a) and the extra help of on-screen text.

Another important issue in relation to frequency of occurrence is the varying size of the effect. An illustration is provided by Rodgers's (2013) study, which was the first to examine the effects of frequency on extensive viewing. In that study participants viewed an episode of a TV show once a week for a period of 10 weeks and took the post-tests one week after the last episode. Rodgers found a significant but small correlation with word frequency on a demanding test format of meaning recognition ($r = .30$) but no significant correlation on a less demanding test ($r = .18$). More generally, in a recent meta-analysis, Uchihara et al. (2019) reported considerable variability in the size of frequency effects on incidental vocabulary learning across studies with an average correlation of .34 (medium effect). The authors concluded that the differences between studies may be due to the intervention of a number of variables such as testing

format, subtitling, learners' vocabulary size, learners' age, range of encounters, and spacing. In relation to spacing, which refers to whether encounters with a lexical item were grouped together (i.e. massed) or spaced out over time (i.e. spaced), it was found that the effect of frequency was stronger in studies conducted in massed conditions and when treatment sessions were completed within a single day. According to Uchihara et al. (2019), the meta-analysis suggests that repeated encounters within a short time span would be more beneficial for incidental word learning (but for evidence in favour of spaced presentations of material in explicit learning contexts, see Ullman and Lovelett, 2018).

Less attention has been given to frequency of occurrence of grammar structures in the input, although it has been assumed that a grammar structure is faster and easier to learn if it appears frequently. This is in line with Goldschneider and DeKeyser's (2001) meta-analysis, which suggested that frequency is one of the main predictors of the order of L2 grammar acquisition. From a constructionist perspective, Ellis and Cadierno (2009) suggested that frequency of exposure and token frequency (how many times a form occurs in the input) are two of the key factors in learning constructions.

Among the scarce existing research, the study conducted by Kusyk and Sockett (2012) from a constructionist perspective focused on the incidental uptake of frequently occurring 4-gram constructions in commonly viewed American TV series. Kusyk and Sockett found a significant correlation between frequency of exposure to TV series and constructions learning: the more time participants spent watching TV series, the better their knowledge of these frequent constructions. In another study, Sockett and Kusyk (2015) showed that frequent viewers of TV series in their original version demonstrated significantly more use of L2 target constructions in their fan-fiction written production than viewers who usually watched versions of the series dubbed into French.

2 On-screen text

The addition of on-screen text in the form of subtitles in L1 or captions in L2 may help learners who are not proficient enough to cope with the fast speech rate and advanced vocabulary presented by authentic audio-visual material (e.g., Guillory, 1998). Winke et al. (2010) explained the usefulness of on-screen text through the fact that it can guide learners' attention to unknown linguistic forms, and promote subsequent noticing and learning through repeated exposure (see also Vanderplank, 2016).

Vocabulary studies have consistently shown that viewing videos with captions is more beneficial for L2 vocabulary learning and comprehension than viewing the same videos without the support of on-screen text (e.g., Jelani and Boers, 2018; Montero Perez et al., 2013). It has been suggested that captions boost vocabulary learning because they help learners segment the speech stream (Charles and Trenkic, 2015), although their effectiveness may depend on the word knowledge aspect that is tested (Montero Perez et al., 2014).

An issue that has deserved some attention concerns the respective benefits of captions and subtitles. Obviously, captions provide learners with more exposure to the target language (e.g. Vanderplank, 2010) which may explain the finding that captions lead to more positive effects on vocabulary learning than subtitles (e.g., Birulés-Muntaner and Soto-Faraco, 2016; Matielo et al., 2015; Peters, 2019). On the other hand, lower-level learners with limited L2 skills may be unable to process and comprehend the L2 audio and text; if a minimum proficiency (and vocabulary size) is necessary to benefit from captions (Neuman and Koskinen, 1992), subtitles might be a better option for supporting beginner learners. However, one particular problem with most existing research is the lack of precise measurement and report of the participants' level of proficiency (Jelani and Boers, 2018; Malone, 2018), making it difficult to compare results across the different studies, or interpret the results of studies with mixed-proficiency participants. The particular word knowledge aspect tested may also have had an influence on research results. In fact, the few studies showing a tendency for subtitle viewing to yield higher gains used meaning recognition or recall tests (Bianchi and Ciabattoni, 2008; Peters et al., 2016), while a stronger link between captions and word-form learning has also emerged (Peters, 2019). Although research aiming at confirming this hypothesis has only found partial support (Peters et al., 2016), it may be suggested that the use of L1-text support is potentially beneficial for word-meaning learning, as indicated by Koolstra and Beentjes (1999) in their pioneer study.

Fewer studies have focused on the role of on-screen text in grammar learning through viewing audio-visual material. Two early studies using reversed subtitling (L1 sound, L2 text) or standard subtitling (L1 text, L2 sound) in short experiments (d'Ydewalle and Van de Poel, 1999; Van Lommel et al., 2006) failed to find significant benefits. It was concluded that grammar may be too complex to learn incidentally (Van Lommel et al., 2006) and a longer period of exposure may be needed than for the acquisition of vocabulary (Kuppens, 2010; Vanderplank, 2016). More recent studies

have analysed the effects of captions and enhanced captions on L2 grammar development. Lee and Révész (2018) found that even though the enhanced captions group outperformed the captions group in the grammatical tasks, both groups significantly improved from the treatment. In a later study by Lee and Révész (2020) including a no captions control group, results indicated that both captioning and textual enhancement were effective. Cintrón-Valentín et al. (2019) compared a no captions to an enhanced captions group and found some evidence of grammar learning from enhanced captioned audio-visual input for half of the target structures. This led the authors to suggest that the structure-specific saliency of a grammar token plays an important role in learning from this type of input (though results were only suggestive because the study lacked a pre-test).

III Learning vocabulary and grammatical constructions through audio-visual input

1 Rationale and aims

The literature review above indicates that extensive viewing of audio-visual input has the potential to optimize L2 learning. As such, it may efficiently support SA students in their pre-departure preparation as well as during SA. This assumption has guided the present paper whose aims are: first, to shed more light on the effects that the repeated encounters with linguistic forms in series episodes may have on their learning; second, to compare the impact of different types of on-screen text and the potential mediating role of proficiency.

With those common aims we conducted two longitudinal studies, focused on vocabulary and grammatical constructions, respectively. Both studies looked into the effects of the number of repeated encounters with the target vocabulary items (TIs) and target grammatical constructions (TCs). Though frequency effects have been investigated in the realm of vocabulary studies, to the best of our knowledge, they have not been previously researched in relation to grammatical constructions. Further, with the exception of Rodgers' (2013) vocabulary study, the effects of frequency of occurrence in viewing studies have been examined from a single session rather than from multiple sessions. The second aim of both studies was to investigate the effects of on-screen text on learning the respective TIs and TCs and the role played by learner proficiency. In the vocabulary study, subtitling and captioning conditions were

compared, whereas in the grammatical constructions study, captioning and no captioning conditions were compared. In both studies English was the target language.

2 Vocabulary study

a Research questions

On the basis of the common aims above, the specific research questions of the vocabulary study were:

1. To what extent is frequency of occurrence of TIs associated to learning outcomes of those items?

2. To what extent do on-screen text (subtitles or captions) and learner proficiency affect L2 vocabulary learning?

b Method

Participants

The initial pool of participants was 53 Catalan-Spanish bilingual adolescents (average age: 13.3) in two 8th grade classes in a public school in Catalonia (Spain). The two intact classes were randomly assigned to one of two learning conditions: subtitling and captioning. Only those participants who had the minimum 85% attendance were included in this study, leaving 39 participants (23 female) for analysis, 17 in one class and 22 in the other. Consent to participate in the study was obtained from their parents through the school.

The Oxford Placement Test (OPT; Allan, 2004) and a vocabulary size test (X_Lex; Meara and Milton, 2003) were administered at the start of the intervention. It was found that learners had a beginner level in English: their average OPT score of 3.03 fell within band 2 (90-104) corresponding to the A1 level of the Common European Framework of Reference (CEFR). Their scores ranged from 56 (Pre-A) to 122 (B1) (see Table 1 below). Their mean vocabulary size was 1,967 words, with scores ranging from 1000 to 3200.

Materials and Target Items

The audio-visual materials were 24 consecutive episodes of an English-spoken TV series that participants were unlikely to have watched (i.e. not yet broadcast on Spanish TV), and that was age appropriate (*Fresh off the Boat*; Khan et al., 2015). There

were a total of 120 TIs (five from each episode), from the first to the nineteenth frequency bands on the BNC/COCA word list (Nation, 2012): 52% belonged to the 1–3 K word families, 21% to the 4–8 K, and 12% to the 9–19 K (15% were off-list). TIs also belonged to different parts of speech, with the majority of them being nouns (60%) and verbs (25%). The TIs were selected from the series according to their frequency of occurrence across the intervention (between 2 and 14 times): 75% occurred between 2 and 5 times, 20% between 6 and 9 times, and 5% between 10 and 14 times. Cognates were excluded from the TIs. See Appendix 1a.

Procedure

The intervention was embedded in a regular English course and administered by the school teachers in collaboration with the researchers. Because of school calendar constraints, the classroom intervention was divided into 3 terms with 8 viewing sessions each (one per week), making a total of 24 episodes and 515 minutes over a period of 8 months. Participants were pre-tested at the beginning and end of each term to assess their knowledge of the corresponding 40 TIs (Figure 1 illustrates one term). The pre- and post-test assessing learners' knowledge of the TIs consisted of listening to aural forms twice and writing them down (written form recall prompted by the aural form of the words), and then providing a translation or a short definition in Catalan or Spanish (meaning recall). This ensured that the tests were congruent with the input-modality (Jelani and Boers, 2018) since written L2 word prompts in the test could have been used to the captions group's advantage. The testing materials were piloted with a comparable group of participants beforehand.

In each session, participants viewed a 20-minute episode (one class with subtitles and one with captions). After viewing each episode, participants completed two tasks: a vocabulary task and a content comprehension task, which were given to better integrate the viewing as a school activity and to encourage learners to pay attention to both vocabulary and content, but they were not corrected in class. The vocabulary task included eight items (both target and non-target items) from the corresponding episode. Participants heard them twice and then had to write them down and select the correct translation amongst the options provided (meaning recognition).[1]

Figure 1.

Preliminary analysis

The comparability of the groups in terms of proficiency was assessed through an independent samples t-test. Results showed that there were no significant differences in the OPT scores for the two groups (t (32) = .128, p = .899). Table 1 displays the descriptive statistics for the participants that completed the proficiency test (five participants did not take the OPT and could not be included in this analysis though they took the pre- and post-tests).

Table 1.

Exploration of the data showed that the proficiency variable was not linearly distributed. Therefore, it was re-categorized into three levels according to the CEFR framework: Pre-A (n = 12), A1 (n = 16) and A2/B1 (n = 6). A two-way group-independence chi-square confirmed that there were no significant differences in the distribution of the two language conditions in the proficiency groups (χ^2 = .283, df = 2, p = .868, with an effect size of V = .09).

Pre- and post-tests were scored dichotomously (0 or 1). For word-form learning, words had to be correctly spelled to be considered correct. For word-meaning learning, translations were scored by two raters (the percent agreement was 95%; disagreement cases were discussed until an agreement was reached). Prior knowledge of the 120 TIs was analysed, and no significant differences were found between the two groups at any of the three pre-testing times, in either form (Term 1: t (37) = -.094, p = .926; Term 2: t (37) = .512, p = .612; Term 3: t (37) = -.223, p = .825) nor meaning (Term 1: t (37) = .913, p = .367; Term 2: t (37) = 1.110, p = .274; Term 3: t (37) = -235, p = .816). See Table 2 for descriptive statistics of scores obtained in the three terms: mean and standard deviation in parenthesis.

Table 2.

The measure of relative gains was calculated per each term and the average relative gains across the three terms was used for the main analysis. A word was considered *learnt* when it was unknown in the pre-test and known in the post-test.

Words known in both pre- and post-test were considered *known* but not *learnt*. Relative gains were calculated at item level following the formula used in previous studies (e.g., Horst et al., 1988; Peters and Webb, 2018):

Relative gains = (number of learnt TIs/(total number of TIs − number of known TIs)) × 100

c *Results*

To address the first research question, Pearson's correlation analyses were run to explore the association between pre-/post-test relative gains in word forms and meanings and token frequency in episodes. Table 3 displays the correlations found for the whole group (N = 39). It can be observed that the number of encounters in the episodes was correlated positively with gains in both word forms and meanings. Table 4 displays the correlations obtained for the subtitles group (n = 17) and the captions group (n = 22) separately.

Table 3.

Table 4.

As can be seen in Table 4, results for the two groups are not identical. For the subtitles group, gains in word meanings correlate significantly with TI frequency. In contrast, for the captions group gains in word forms correlate significantly with TI frequency.

The second research question concerned the extent to which the language of the on-screen text (L1 / L2) and learners' proficiency have an effect on vocabulary learning through viewing the same videos. A set of generalized linear models (GLMs) were run to evaluate the influence of these factors on the two vocabulary outcome measures: word-form learning and word-meaning learning.

Word-form learning

A GLM was first calculated with the percentage of word-forms learnt as the dependent variable, and language condition (subtitling or captioning) and proficiency

(Pre-A, A1 or A2/B1) as fixed effects, including the interaction between language condition and proficiency. The results showed that the interaction was not significant and a new model was run. Table 5 below presents the final model, and Table 6 reports the pairwise contrasts between categorical factors.

Table 5.

Table 6.

The model revealed a main effect of proficiency, but no main effect of language of the on-screen text. Results showed that the most proficient students (A2/B1 level) scored 6.53% higher than A1-level students and 11.50% higher than Pre-A-level students. Pairwise contrasts revealed that differences were only significant between Pre-A and A2/B1 ($p = .003$).

Word-meaning learning

A GLM was calculated with word-meaning gains (relative gains in percentage for meaning recall across the intervention) as the dependent variable, and language (subtitles or captions) and proficiency (Pre-A, A1 or A2/B1) as fixed effects. The interaction between language condition and proficiency was also included. The results again showed that the interaction was not significant and a new model was run. Table 7 presents the fitted model and Table 8 reports the pairwise contrasts between categorical factors.

Table 7.

Table 8.

Similar to gains in form, the model showed a main effect of proficiency for word-meaning learning, but no main effect of language of the on-screen text. Results showed that gains in meaning depended significantly on participants' proficiency level ($p = .001$), with the most proficient students scoring 11.66% higher than the least proficient students. Pairwise contrasts revealed that differences were significant between A2/B1 and both Pre-A ($p = .001$) and A1 ($p = .038$), but not between A1 and Pre-A level ($p = .160$).

3   Grammatical constructions study

a   Research questions

The specific research questions of the grammar study were:

1. To what extent is frequency of occurrence of TCs associated to learning outcomes of those constructions?

2. To what extent do on-screen text (captions or no captions) and learner proficiency affect L2 grammatical constructions learning?

b   Method

Participants

The initial pool comprised 90 undergraduates at a public university in Catalonia, but 19 were excluded from the data analysis because they attended fewer than 8 of the 10 sessions. Thus, the final number was 69 participants (41 female). They were Catalan-Spanish bilinguals ranging in age from 17 to 32 (average: 19.3). The participants had been randomly distributed in two classes prior to the intervention and the classes were randomly assigned to one of two conditions: captioning (n=39) and no captioning (n=30). Consent was obtained from the students themselves (credits were allocated for their participation). The results of the OPT administered at the beginning of the intervention showed great variability. Their mean score of 139.55 fell within OPT band 5 (135-149), corresponding to the B2 level of the CEFR; however, scores ranged from 92 to 183, that is, CEFR levels A1 to C2 (see Table 9 below).

Materials and Target Constructions

The participants viewed ten consecutive episodes of a TV series (*The Good Place*; Schur, 2016) that they were unlikely to have watched (i.e. not yet broadcast on Spanish TV). From the scripts of the series, a total of 16 abstract or grammatical constructions were selected. These are constructions which can allow for rule abstraction (Tomasello, 2005), created through pattern-finding (i.e., categorization, analogy, and distributional analysis), which is considered the central cognitive construct in the so-called usage-based approach to the acquisition of grammar (Goldberg, 2006; see Tomasello, 2009:70). A correlational analysis indicated a significant relationship ($r = .611$, $p = .020$) between the frequency of encounters of the selected TCs and their frequency in the TV Corpus (Davies, 2019) – a part of The Corpus of Contemporary

American English (Davies, 2008) containing informal language from TV shows (1950-2018). Thus, the TCs in the study can be claimed to be ecologically valid and representative of the language used in English media sources.

Only constructions which occurred at least 3 times in the selected episodes were included (see Appendix 1b). Their frequency of occurrence across the intervention was: 44% between 3 and 10 times, 31% between 11 and 20 times, 19% between 21 and 23 times, and 6% more than 23 times. The testing materials were piloted with a comparable group of participants beforehand.

Procedure

The intervention was embedded in a regular English course and administered by their teacher in collaboration with the researchers. The testing materials for the pre-/post-tests of the TCs consisted of productive grammar exercises (sentence transformation, fill-the-gap, and complete the gap with a correct form of a given word). Participants were not informed of the nature of the experiment and the teacher did not provide any extra practice on grammar, neither were students notified beforehand about the upcoming tests.

The two classes viewed the same ten episodes, one with captions and one without captions, over five weeks; two different episodes per week on two different days (see Figure 2). Each episode's length was approximately 21 minutes (totaling 227 minutes of audio-visual input). Both classes completed the same post-viewing tasks that were given to better integrate the viewing as a classroom activity and to encourage learners to pay attention to language and content. Each week, at the end of the first viewing session, participants completed a recognition task in which they had to decide whether they had heard (and read in the captioned condition) a set of distractor non-target constructions, and a comprehension task. Two days after, the participants watched the next episode and completed those tasks and an additional one: a grammar task that elicited a maximum of three of the TCs from the preceding two episodes.[2]

Figure 2.

Preliminary analysis

To test whether there were any differences in proficiency between the captions and no captions groups before the intervention (see Table 9), an independent samples t-

test was conducted. The results showed that both groups were comparable (t (69) =. 547, p = .372).

Table 9.

The variable proficiency was not linearly distributed so it was recoded into three different groups following the OPT scoring procedure and CEFR framework: A1-A2 (n = 14), B1-B1 (n = 32), and C1-C2 (n = 23). There were no significant differences in distribution between the two conditions (χ^2 = .420, df = 2, p = .811, with an effect size of V = .08).

Pre- and post-tests were scored dichotomously (0 or 1). There were no significant differences between pre-test scores of the two groups (t(69) =. 064, p = .961). Table 10 displays the descriptive statistics (mean and standard deviation) of the scores in the pre-test and post-test.

Table 10.

c   **Results**

Table 11 displays the Pearson's correlations of the participants' relative gains in the TCs (calculated using the same formula as in the vocabulary study) with their token frequency in the episodes for all the participants together (N = 69) and for the two groups separately (n = 39 and n = 30). The correlation is significant for the whole group and for the group who watched the episodes without captions. However, for the group who viewed the episodes with captions the correlation is lower and marginal.

Table 11.

To explore the influence of on-screen text and proficiency on these participants' learning of grammatical constructions, a GLM was run. The post-test scores on the 16 target constructions was the dependent variable; group (captions vs. no captions) and proficiency (elementary, intermediate or advanced) were included as main factors. The interaction between group and proficiency was also introduced in the model. Pre-test scores were included as the covariate to control for learners' previous knowledge of

grammatical constructions. A first model showed that the interaction between group and proficiency was not significant and a second model was run without it. The final model (see Table 12) showed a main effect for group, where the captions group outperformed the no captions group, and a main effect of pre-test scores and of proficiency. Table 13 reports the pairwise contrast between categorical factors.

Table 12.
Table 13.

The pairwise comparisons between the proficiency groups showed that learning outcomes did not have a linear relationship with proficiency, as the advanced group did not learn more than the intermediate group through watching the episodes. The intermediate proficiency group had significantly higher gains than the elementary group ($p < .001$). Yet, no significant difference was found in the comparisons either between the elementary and the advanced groups or between the intermediate and the advanced groups.

4   General discussion

This paper has presented two longitudinal studies, one focused on vocabulary and one on grammar, examining the potential influence of two aspects of audio-visual input from TV series: repeated encounters with linguistic forms and on-screen text support. The results of both studies shed light on the conditions under which frequency effects are stronger and on the role played by the various modes of on-screen text in relation to learners' level of proficiency in language learning.

a   Frequency effects

The results of both studies concurred in showing the impact of repeated exposures on learning, but the size of the frequency effect varied in interesting ways. The vocabulary study focused on the frequency of encounters of the TIs and its association with vocabulary gains by adolescents with an elementary level of proficiency. The results of the correlational analyses showed a significant, if modest, relationship: token frequency was positively and significantly correlated with gains in word-form and word-meaning learning, in agreement with previous research findings in

the area of audio-visual input (Peters, 2019; Peters et al., 2016; Peters and Webb, 2018; Rodgers, 2013). A more interesting result emerged when running the correlations separately for the subtitles and captions groups: the number of encounters was significantly associated with word-meaning gains for the subtitles group and with word-form gains for the captions group. This difference was nonconflicting: the subtitles group was provided with the meaning in L1 and repeated encounters enhanced learning; in contrast, the captions group did not have direct access to the meaning and frequency was not a contributing factor in their learning of English word meanings. This learning has been reported to come at a more advanced stage (Godfroid et al., 2018) and is one of the most difficult aspects of word knowledge to acquire incidentally (Van Zeeland and Schmitt, 2013). On the other hand, the captions group were able to read the word forms in English and their learning was improved with higher frequency; in the subtitles group participants did not see the English word forms during the viewing, and so their frequency in the videos was less relevant for word-form learning (note that a learner may not have fully acquired the correct word form – for example, they may misspell it, and still attach the correct meaning to it).

An intriguing finding of the present study was that the correlations, even when significant, were smaller than in most previous vocabulary studies. The combination of on-screen text and visual images may have attenuated frequency effects, as Uchihara et al. (2019) noted in their meta-analysis. Specifically for this study, a plausible explanation may lie in differences in spacing, as predicted by Webb (2014) and observed by Uchihara et al. (2019), where higher correlations were observed in massed treatment conditions (one-day treatment exposure) than in spaced treatment conditions (multiple treatment exposures over an extended period of time). In other words, the smaller frequency effects recorded in this vocabulary study may be characteristic of extensive viewing studies that employ authentic material in which word repetitions are distributed sparsely over several sessions separated from each other by days or weeks, and in which there may be substantial decay in knowledge from the first weeks until the post-test (Rodgers, 2013). On that note, the size of the correlations found in the current study is similar, for example, to that found by Webb and Chang (2015) from an extensive reading program with audio support over 13 weeks. Another plausible partial explanation for the weak correlations in this study is participants' age. No previous research has explored age effects on the relationship between frequency of encounters and learning, but Uchihara et al's (2019) meta-analysis found a larger frequency effect

for university learners than for secondary school learners. Similarly, in a meta-regression examining the effectiveness of incidental L2 word learning from spoken input, De Vos et al. (2018) found significant effects for age, which the authors attributed to a number of factors: one of these factors was the longer experience with the L2 of the university students which helped them to use strategies to derive the meanings of unknown words more effectively.

The grammar study investigated the association between frequency of encounters of the TCs and the learning gains of those constructions by two groups of upper-intermediate university students, one of which viewed the videos with captions and the other without. The results of the correlational analysis showed a significant correlation when the two groups were considered together, which corroborates previous findings of research into construction learning (e.g., Kusyk and Sockett, 2012). However, when the correlations were run separately for the two groups, the correlation was confirmed for the no captions group but became marginal for the captions group. This result seems to indicate that repetition of the target constructions was more beneficial for finding the construction patterns for the group that viewed the episodes without captions (the more challenging condition) than for the captions group. To our knowledge, no previous study has examined the effects of construction repetition in captioned and non-captioned audio-visual input, but on-screen text has been observed to attenuate frequency effects in vocabulary studies (as noted above). Our finding suggests that a frequency effect emerges more clearly when learners do not have other types of more effective support, i.e., captions. Similar findings of frequency effects arising in a more challenging condition are observed in the realm of vocabulary. For example, larger effects have been found with more demanding measures, e.g., a recall format compared to a recognition format (Uchihara et al., 2019) and more demanding target words, e.g., non-cognate words compared to cognate words (Horslund, 2019).

The size of the frequency effect also differs across the two studies, being much larger in the grammar study than in the vocabulary study. While the current data do not allow for a direct comparison between frequency effects in grammar and vocabulary learning, two reasons may account for some of the difference. The first is learners' age, which favours university learners over secondary school learners, as seen above. The second may lie in the respective spacing conditions - one episode a week for the adolescent learners compared to two episodes a week for the young adult learners - as well as in the interval between TIs and TCs encounters and post-test, potentially

extending over 8 weeks and 5 weeks, respectively (given that very long gaps may result in diminished retention; see Cepeda et al., 2008).

b   ***On-screen text and proficiency***

The second research question in both studies addressed the role played by on-screen text and learners' proficiency, by comparing subtitling and captioning in the vocabulary study, and captioning and no captioning in the grammar study. In the vocabulary study the analyses were conducted separately for word-form and word-meaning recall. For both aspects of vocabulary, the analysis showed a main effect of proficiency, but no main effect of language, and no interaction between language condition and proficiency. Results showed that more advanced learners obtained higher gains, in line with some previous research (Chen et al., 2018). For word form, the differences were significant between A2/B1and Pre-A, and for word meaning, they were significant between A2/B1 and both Pre-A and A1, but not between A1 and Pre-A levels. The lack of a statistical difference between the subtitles and captions groups seems to indicate that their level of proficiency was too low to derive the benefits shown by previous research with more advanced viewers of captioned videos. Neither was the captions group significantly outperformed by the subtitles group, which suggests that at this age and proficiency level both on-screen text types may be similar in terms of their adequacy for vocabulary learning. This result corroborates those of some previous studies but diverges from others which have shown a learning advantage for either captions (the majority) or subtitles. For example, Bravo's study (2008) did not find a significant difference with young participants at A2/B1 level (in an intervention that included attention to form). In contrast, with slightly older (aged 15-18) and more proficient learners, the studies by Naghizadeh and Darabi (2015) and by Peters et al. (2016) found that the captions groups performed significantly better in vocabulary learning than the subtitles groups. As suggested in Pujadas and Muñoz (2019), there may be a proficiency threshold: that is, the more proficient the learner (and usually, the older as well), the more they benefit from captioning rather than subtitling.

Another factor that may have played a role is the word knowledge aspect examined, as the test used in the current study, involving knowledge of both form and meaning, may have neutralized the potential advantage of the captioning condition with these learners. Indeed, as seen in the analysis of frequency effects above, a link emerged between captions and word form and between L1 subtitles and word meaning.

In the grammar study the analysis showed that the captions group outperformed the no captions group. This corroborates results from previous research showing that captions can facilitate the acquisition of L2 grammatical constructions (Lee and Révész, 2020), and extends the results obtained for L2 listening and vocabulary (e.g., Montero Perez et al., 2013) to grammatical knowledge. The results also show that learning outcomes did not have a linear relationship with proficiency: the intermediate group had higher gains than the elementary proficiency group but no difference in gains was shown between the other groups. This may be explained by the nature of the input in the TV series, which may have contained a lower quantity of novel constructions for the advanced group than for less advanced learners, thus limiting the potential room for development, and may therefore have been especially well attuned to intermediate learners.

The combined results of these analyses highlight the crucial role played by proficiency and provide a sound explanation for the previous mixed results that favoured one type of on-screen text over others without having taken due account of participants' proficiency level. The current results are in agreement with a suggestion by Gass et al. (2019: 87) of a zone within which captions –and subtitles, we might add – "are most useful, perhaps centered where the content difficulty level is not too far above or below a learner's ability level."

5   Conclusion

The present investigation makes several relevant contributions to the increasing body of research into the benefits of audio-visual input and multimodal input in general (e.g., Mayer, 2014; Paivio, 1986) for L2 acquisition. Firstly, it provides data from extensive treatments, in which repetitions are spaced, in contrast to the majority of studies that investigate learning from a single viewing session; the study, thus, highlights the moderating role of spacing in frequency effects. The results also extend the frequency effects often found for item-based lexical items to the area of grammatical constructions, an under-researched area in relation to audio-visual input, and suggest that learners are able to find the patterns of frequent constructions in this type of input. Furthermore, it shows that for both vocabulary and grammar the size of the effect may depend on the type of support provided by on-screen text. In this way, the present results add nuance to the expected frequency effects and contribute to our knowledge about important variables moderating the relationship between the frequency of

encounters and vocabulary learning. In addition, our knowledge base of the role played by on-screen text has been enlarged with the inclusion of its effects on grammar and the recognition of the important role played by learner proficiency, a finding which may help explain the mixed results obtained in previous research.

Inevitably, this study is not without its limitations. Extensive classroom studies are more ecologically valid but less tightly controlled, and the possibilities of learners encountering the target items outside the classroom over the intervention period are greater; in addition, the inclusion of the post-viewing tasks to better integrate the interventions in a regular classroom may have enhanced certain target forms in a non-perceptible way. Another limitation is the potential influence of the difficulty level of the series, since series with lower or higher lexical coverage may have yielded different results with respect to on-screen text and proficiency. Finally, though comparisons were made between the two studies when they highlighted certain interesting aspects, this was not intended as a comparative study (a design that would have required parallel conditions).

IV   Lessons for SA research and pedagogical implications

The contributions of these findings are relevant to input-related concerns in the area of SA in various ways. They have identified incidental learning through exposure to multimodal input, which helps explain the effectiveness of SA for vocabulary and grammar learning in multimodal real-life situations (e.g. pictures + text, gesture + voice) in which sojourners participate. More specifically, the results have shed light on important mediating factors. Repeated exposures that characterize social routines embedded within SA sojourners' daily activities may make language elements more salient and may facilitate learning. Frequency effects are stronger in the absence of other types of support such as on-screen text, as in real-life interaction. The current results also highlight the important role of learners' age and proficiency, showing how they interact with the characteristics of the input, specifically with the support provided by different types of on-screen text. Attention to these factors may partly explain the inter-individual variability found in SA research.

These results have implications for teachers and learners as well. They are relevant to teachers in pre-departure preparation programs giving orientations to learners about how to maximize their preparation and their stays abroad in accordance with their proficiency level. They highlight the value of multimodal comprehension

practice using authentic speech. During SA, extensive viewing of audio-visual material may be one type of support that complements natural interaction with other target language speakers but allowing learners to regulate the ceaseless flow of L2 input that may characterize SA. Furthermore, viewing is free of the pressure of social interaction which may become overwhelming at times (Muñoz, 2017) and gives learners room for self-regulation and assessment of their learning process.

---

1 All the TIs appeared only once in the post-viewing tasks. The aim of this task was to compare these groups with two other groups that had been pre-taught the TIs (see Pujadas and Muñoz, 2019). To account for the potential learning effects of this task, the scores in the pre-test and in the post-viewing task were compared for a sample (50% of tasks in a term). A Wilcoxon Signed Rank test indicated significant gains in both word forms and meanings through viewing the videos before completing the post-viewing task.

2 The aim of this task was to examine recency effects (not analysed in this study). To account for the potential learning effects of this task, a set of comparisons were made between the scores in the pre-test and each one of the five grammar tasks in the intervention. A Wilcoxon Signed Rank test indicated that all participants significantly improved their knowledge of L2 constructions through viewing the videos before completing the post-viewing task.

References

Allan, D (2004) *Oxford Placement Test 1*. Oxford: Oxford University Press.

Bianchi, F, Ciabattoni, T (2008) Captions and subtitles in EFL learning: an investigative study in a comprehensive computer environment. In: Baldry, A, Pavesi, M,Taylor Torsello, C (eds) *From didactas to ecolingua.* Trieste: Edizioni Università di Trieste, pp. 69–90.

Birulés-Muntané, J, Soto-Faraco, S (2016) Watching subtitled films can help learning foreign languages. *PLOS ONE* 11 (6): e0158409.

Bravo, MCC (2008) *Putting the reader in the picture. Screen translation and foreign-language learning.* PhD Thesis, University Rovira i Virgili, Spain, and University of Algarve, Portugal.

Brecht, RD, Davidson, D, Ginsberg, R (1990) The empirical study of proficiency gain in study abroad environments of American students of Russian. In: Davidson, D

(ed.) *American contributions to the VII International Congress of MAPRIAL.* Washington, DC: American Council of Teachers of Russian, pp.123–152.

Brecht, RD, Robinson, JL (1993) *Qualitative analysis of second language acquisition in study abroad: The ACTR / NFLC project*. National Foreign Language Center.

Brown, R, Waring, R, Donkaewbua, S (2008) Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a foreign language* 20(2): 136–163.

Cepeda, NJ, Vul, E, Rohrer, D, Wixted, JT, Pashler, H (2008) Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science* 19: 1095–1102.

Charles, T, Trenkic, D (2015) Speech segmentation in a second language: The role of bimodal input. In: Gambier, Y, Caimi, A, Mariotti, C (eds) *Subtitles and language learning: Principles, strategies and practical experiences*. Bern: Peter Lang, pp.173–198.

Chen, C, Truscott, J (2010) The effects of repetition and L1 lexicalisation on incidental vocabulary acquisition. *Applied Linguistics* 31: 693–713.

Chen, Y, Liu, Y, Todd, AG (2018) Transient but effective? Captioning and adolescent EFL learners' spoken vocabulary acquisition. *English Teaching & Learning* 42 (1): 25–56.

Cintrón-Valentín, M, Garcia-Amaya, L, Ellis, N (2019) Captioning and grammar learning in the L2 Spanish classroom. *The Language Learning Journal* 47(4): 1–21.

Collins, L, Muñoz, C (2016) The Foreign Language Classroom: Current Perspectives and Future Considerations. *The Modern Language Journal* 100 (S1): 133–147.

Cubillos, J, Chieffo, L, Fan, C (2008) The impact of short-term study abroad programs on L2 listening comprehension skills. *Foreign Language Annals* 41: 157–185.

Davies, M (2008-) *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. Available at: https://www.english-corpora.org/coca/. (accessed 1 August 2019).

Davies, M (2019-) *The TV Corpus: 325 million words, 1950-2018*. Available at: https://www.english-corpora.org/tv/ (accessed 1 August 2019).

DeKeyser, RM (2007) Study abroad as foreign language practice. In: DeKeyser, R (ed.) *Practice in a Second Language. Perspectives from Applied Linguistics and Cognitive Psychology*. NY: Cambridge University Press, pp.208–226.

DeKeyser, RM (2010) Monitoring processes in Spanish as a second language during a study abroad program. *Foreign Language Annals,* 43(1): 80–92.

De Vos, JF, Schriefers, H, Nivard, MG, Lemhöfera, K (2018) A Meta-Analysis and Meta-Regression of Incidental Second Language Word Learning from Spoken Input. *Language Learning* 68(4): 906–941.

D'Ydewalle, G, Van de Poel, M (1999) Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research* 28(3): 227–244.

Ellis, N, Cadierno, T (2009) Constructing a Second Language Introduction to the Special Section. *Annual Review of Cognitive Linguistics* 7: 111–139.

Faretta-Stutenberg, M, Morgan-Short, K (2018) Contributions of initial proficiency and language use to second-language development during Study Abroad. Behavioral and event-related potential evidence. In: Sanz, C, Morales-Front, A (eds) *The Routledge Handbook of Study Abroad Research and Practice*. Abington: Routledge, pp. 421–435.

Gass, S, Winke, P, Isbell, DR, Ahn, J (2019) How captions help people learn languages: A working-memory, eye-tracking study. *Language Learning & Technology* 23(2): 84–104.

Goldberg, A (2006) *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Goldberg, A, Casenhiser, D (2006b) English Constructions. In: McMahon A and Aarts B (eds) *The Handbook of English Linguistics.* Blackwell Publishers, pp. 343–355.

Godfroid, A, Ahn, J, Choi, I, Ballard, L, Cui, Y, Johnston, S, Lee, S, Sarkar, A, Yoon, HJ (2018) Incidental vocabulary learning in a natural reading context: an eye-tracking study. *Bilingualism: Language and Cognition* 21(3): 563–584.

Guillory, H (1998) The effects of keyword captions to authentic French video on learner comprehension. *CALICO Journal* 15(1-3): 89–108.

Gullberg, M, Roberts, L, Dimroth, C (2010) What word-level knowledge can adult learners acquire after minimal exposure to a new language? *International Revue of Applied Linguistics* 50: 239–276.

Hashimoto, BJ, Egbert, J (2019) More than frequency? Exploring predictors of word difficulty for second language learners. *Language Learning* 69 (4): 839 – 872. .

Horslund, CS (2019) English vocabulary in L1 Danish and L1 Finnish learners: Vocabulary sizes, word frequency effect, and cognate facilitation. In:

Christensen, KR, Jørgensen, H, Wood, JL (eds) *The Sign of the V – Papers in Honour of Sten Vikner*. Dept. of English, School of Communication & Culture, Aarhus University, pp. 267–283.

Horst, M, Cobb, T, Meara, P (1998) Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language* 11(2): 207–223.

Jelani, NAM, Boers, F (2018) Examining Incidental Vocabulary Acquisition from Captioned Video. Approaches to Learning, Testing, and Researching L2 Vocabulary. *ITL - International Journal of Applied Linguistics* 169(1): 169–90.

Khan, N, Kasdar, J, Melvin, M, Blomquist, R, Huang, McEwen, J (2015) *Fresh off the Boat* [TV series] Los Angeles, CA: ABC.

Koolstra, CM, Beentjes, JWJ (1999) Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational technology research and development* 47: 51–60.

Kuppens, AH (2010) Incidental foreign language acquisition from media exposure. *Learning, Media and Technology* 35: 65–85.

Kusyk, M, Sockett, G (2012) From informal resource usage to incidental language acquisition: language uptake from online television viewing in English. *Asp*, *la revue dun GERAS* 62: 45–65.

Lee, M, Révész, A (2018) Promoting Grammatical Development Through Textually Enhanced Captions: An Eye-Tracking Study. *The Modern Language Journal* 102: 557–577.

Lee, M, Révész, A (2020). Promoting Grammatical Development through Captions and Textual Enhancement in Multimodal Input-based Tasks. *Studies in Second Language Acquisition.* Epub ahead of print 22 April 2020. DOI: 10.1017/S0272263120000108.

Llanes, À (2011) The many faces of study abroad: an update on the research on L2 gains emerged during a study abroad experience. *International Journal of Multilingualism* 8: 189–215

Malone, J (2018) Incidental Vocabulary Learning in SLA. *Studies in Second Language Acquisition* 40(3): 1–25.

Matielo, R, Collet, T, D'Ely, RCSF (2013) The effects of interlingual and intralingual subtitles on vocabulary learning by Brazilian EFL learners: an exploratory study.

Intercâmbio. Revista do Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem 27: 83–99.

Matielo, R, D'Ely, RCSF, Baretta, L (2015) The effects of interlingual and intralingual subtitles on second language learning/acquisition: a state-of-the-art review. *Trabalhos em Linguística Aplicada* 54: 161–182.

Mayer, R (2014) *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

Meara, P, Milton, J (2003) *X_Lex: the Swansea levels test*. Berkshire: Express Publishing.

Montero Perez, M, Van Den Noortgate,W, Desmet, P (2013) Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System* 41: 720–739.

Montero Perez, M, Peters, E, Clarebout, G, Desmet, P (2014) Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology* 18(1): 118–41.

Muñoz, C (2017) Tracing trajectories of young learners: Ten years of school English learning. *Annual Review of Applied Linguistics* 37: 168–184.

Naghizadeh, M, Darabi, T (2015) The Impact of Bimodal, Persian and No-Subtitle Movies on Iranian EFL Learners' L2 Vocabulary Learning. *Journal of Applied Linguistics and Language Research* 2(2): 66–79.

Nation, P (2007) The four strands. International Journal of Innovation in Language *Learning and Teaching* 1(1): 2–13.

Nation, P (2012) The BNC/COCA Word Family Lists. Document bundled with Range program with BNC/COCA lists, 25.

Nation, P (2015) Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language* 27(1): 136–145.

Neuman, S, Koskinen, P (1992) Captioned television as comprehensible Input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly* 27(1): 94–106.

Paige, RM, Cohen, AD, Kappler, B, Chi, JC, Lassegard, JP (2002) *Maximizing study abroad. A students' guide to strategies for language and culture learning and use.* Minneapolis: University of Minnesota Center for Advanced Research on Language Acquisition.

Paivio, A (1986) *Mental representations: A dual coding approach*. Oxford: Oxford University Press.

Pattemore, A, Muñoz, C (2020) Learning L2 constructions from captioned audio-visual exposure: The effect of learner-related factors. *System* 93: https://doi.org/10.1016/j.system.2020.102303

Pellicer-Sánchez, A (2016) Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition* 38(1): 97–130.

Peters, E (2019) The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *TESOL Quarterly* 53(4): 1008–1032.

Peters, E, Heynen, E, Puimège, E (2016) Learning vocabulary through audiovisual input: the differential effect of L1 subtitles and captions. *System* 63: 134–48.

Peters, E, Webb, S (2018) Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition* 40(3): 551–77.

Pujadas, G, Muñoz, C (2019) Extensive viewing of captioned and subtitled TV series: A study of L2 vocabulary learning by adolescents. *The Language Learning Journal* 47: 479–496.

Rivers, WP (1998) Is being there enough? The effects of homestay placements on language gains during study abroad. *Foreign Language Annals* 31(4): 492–500.

Rodgers, M (2013) English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions. PhD Thesis, Victoria University of Wellington, New Zealand.

Rodgers, M (2018) The images in television programs and the potential for learning unknown words. The relationship between on-screen imagery and vocabulary. *ITL - International Journal of Applied Linguistics* 169(1): 191–211.

Rodgers, M, Webb, S (2011) Narrow viewing: The vocabulary in related television programs. *TESOL Quarterly* 45(4): 689–717.

Rodgers, M, Webb, S (2019) Incidental vocabulary learning through viewing television. *ITL International Journal of Applied Linguistics*. Epub ahead of print 25 June 2019. DOI: 10.1075/itl.18034.rod.

Sanz, C, Morales-Front, A (2018) *The Routledge Handbook of Study Abroad Research and Practice*. Abington: Routledge.

Schmitt, N, Carter, R (2000) The lexical advantages of narrow reading for second language learners. *TESOL Journal* 9: 4–9.

Segalowitz, N, Freed, BF (2004) Context, contact, and cognition in oral fluency acquisition. Learning Spanish in At Home and Study Abroad contexts. *Studies in Second Language Acquisition* 26: 173–199.

Shur, M (2016) *The Good Place* [TV series]. United States: Fremulon.

Sockett, G, Kusyk, M (2015) Online informal learning of English: frequency effects in the uptake of chunks of language from participation in web-based activities. In: Cadierno, T, Eskildsen, SW (eds) *Usage-based perspectives on second language acquisition.* Berlin: De Gruyter Mouton, pp.153–177.

Sydorenko, T (2010) Modality of input and vocabulary acquisition. *Language Learning & Technology* 14: 50–73.

Tomasello, M (2005) *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Tomasello, M (2009) The usage-based theory of language acquisition. In Bavin, E (ed.) *The Cambridge Handbook of Child Language (Cambridge Handbooks in Language and Linguistics)*. Cambridge: Cambridge University Press, pp. 69–88.

Tschinner, E (2007) The development of oral proficiency in a four-week intensive immersion program in Germany. *Die Unterrichts praxis/Teaching German,* 40(2): 111–117.

Tullock, B, Ortega, L (2017) Fluency and multilingualism in study abroad: Lessons from a scoping review. *System* 71: 7–21.

Uchihara, T, Webb, S, Yanagisawa, A (2019) The Effects of Repetition on Incidental Vocabulary Learning: A Meta-Analysis of Correlational Studies. *Language Learning* 69(3): 559–99.

Ullman, MT, Lovelett, JT (2018) Implications of the declarative/procedural model for improving second language learning: The role of memory-enhancing techniques. *Second Language Research* 34(1): 39–65.

Van Lommel, S, Laenen, A, D'Ydewalle, G (2006) Foreign-grammar acquisition while watching subtitled television programmes. *British Journal of Educational Psychology* 76(2): 243–258.

Van Zeeland, H, Schmitt, N (2012) Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics* 34(4): 457–79.

Van Zeeland, H, Schmitt, N (2013) Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System* 41(3): 609–624.

Vandergrift, L (2007) Recent developments in second and foreign language listening comprehension research. *Language Teaching* 40: 191–210.

Vanderplank, R (2010) Déjàvu? A decade of research on language laboratories, television and video in language learning. *Language Teaching* 43(1): 1–37.

Vanderplank, R (2016) 'Effects of' and 'effects with' captions: how exactly does watching a TV programme with same language subtitles make a difference to language learners? *Language Teaching* 49(2): 235–50.

Vidal, K (2011) A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning* 61(1): 219–258.

Webb, S (2014) Repetition in incidental vocabulary learning. In: Chapelle, CA (ed.) *The Encyclopedia of Applied Linguistics*. Hoboken, NJ: Wiley, pp. 1–6.

Webb, S (2015) Extensive viewing: Language learning through watching television. In: Nunan, D, Richards, J (eds) *Language learning beyond the classroom.* Abingdon: Routledge, pp.175-184.

Webb, S, Chang, A (2015) Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research* 19: 667–686.

Webb, S, Rodgers, M (2009a) Vocabulary demands of television programs. *Language Learning* 59(2): 335–66.

Webb, S, Rodgers, M (2009b) The lexical coverage of movies. *Applied Linguistics* 30(3): 407–27.

Winke, P, Gass, S, Sydorenko, T (2010) The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology* 14: 65–86.

Wisnieska, N, Mora, JC (2020). Can captioned video benefit second language pronunciation? *Studies in Second Language Acquisition* 42(3): 599–624. DOI: 10.1017/S0272263120000029.

Appendix 1a

Target vocabulary items with frequency of occurrence

	Target items	N (%)
2	to ace, buckle, cabbage, curly, to gamble, to gloat, haircut, hairdryer, hedgehog, kilt, to loiter, orchids, prank, principal, straight, stuffed animal, threat to tease, to rush, trap to trust	21 (17.5%)
3	binder, cheers, Christmas Eve, cushy, drill, dump, fake, to fit in, gross, guest, janitor, jukebox, knockoff, to hand out, to hang out, necklace, to mess up, nightmare, oyster, to pick on, teamwork, to trade, whale	23 (19.2%)
4	to bounce, cranberry, fridge, to ground, to hire, Jewish, nerd, oven, real, rug, shield, sleepover, van	13 (10.8%)
5	bill, bow, bunk, to cash, to cut off, fortune teller, friendship, to join, proud, to quit, raw, ride, to trip, wealth	14 (11.7%)
6	bakery, to bet, carpool, chess, to dare, dull, franchise, to figure out, napkin, neighbourhood, rib, to struggle, wall	13 (10.8%)
7	coach, fireworks, floor mat, to handle, to make sense, realtor, review, scary, turkey	9 (7.5%)
8	chickenpox, crouton, flight attendant, noodle, ride, to share, tax	7 (5.8%)
9	grade, hen, piccolo, weird	4 (3.3%)
10	AC, billboard, costume, mall,	4 (3.3%)
11-15	to date, deal, fee, to fire, tough, jealous, play, hell, envelope, enough	10 (8.4%)
16-20	cool, to move	2 (1.7%)

Total frequency of occurrence from pre- to post-test

Appendix 1b

Target constructions with frequency of occurrence

Construction	Examples from *The Good Place*	Frequency in *The Good Place*	Frequency in The TV Corpus
Verb phrase	*Let's do* it again	55	246479
	I *did* wash the dishes!	23	95969
	I knew you *were going to come*	22	3644
	Let me show you	21	126232
	You *want her to* leave	19	128926
	You'*re* not *supposed to* see it	18	83675
	Okay, you're right. I *don't belong here*	17	1501
	I promised that I *would help*	15	4371
	I'm better than I *used to* be	7	23329
	I can't advise you to be dishonest *either*	5	7520
	I *will be joining* them	4	6475
	I wish I had your wisdom	4	2121
	Architects *aren't allowed to* own	3	10158
Passive	Every detail *has been designed*	14	13588
	They *are being used*	3	22897
Comparative Correlative	*The sooner* she's gone, *the better*	4	188

Adapted from Goldberg and Casenhiser (2006)