

Captioning and grammar learning in the L2 Spanish classroom

Myrna Cintrón-Valentín, Lorenzo García-Amaya, and Nick C. Ellis

QUERY SHEET

This page lists questions we have about your paper. The numbers displayed at left are hyperlinked to the location of the query in your paper.

The title and author names are listed on this sheet as they will be published, both on your paper and on the Table of Contents. Please review and ensure the information is correct and advise us if any changes need to be made. In addition, please review your paper as a whole for typographical and essential corrections.

Your PDF proof has been enabled so that you can comment on the proof directly using Adobe Acrobat. For further information on marking corrections using Acrobat, please visit <http://journalauthors.tandf.co.uk/production/acrobat.asp>; <https://authorservices.taylorandfrancis.com/how-to-correct-proofs-with-adobe/>

The CrossRef database (www.crossref.org/) has been used to validate the references. Changes resulting from mismatches are tracked in red font.

AUTHOR QUERIES

QUERY NO.	QUERY DETAILS
Q1	Please note that the ORCID has/have been created from information provided through CATS. Please correct if this is inaccurate.
Q2	The abstract is currently too long. Please edit the abstract down to no more than 200 words.
Q3	Funding details (National Science Foundation) have been taken from information supplied with your manuscript submission and checked against the Open Funder Registry and we found a partial match “National Science Foundation”. Please check and resupply the funding details.
Q4	Funding details have been taken from information supplied with your manuscript submission. Please correct if this is inaccurate.
Q5	A Quotation mark seems to be missing following “ the stimulus through the subject”. Please indicate where it should be placed.
Q6	The disclosure statement has been inserted. Please correct if this is inaccurate.
Q7	Please provide missing city for the reference "Blake 2013" references list entry.



Captioning and grammar learning in the L2 Spanish classroom

Q1 Myrna Cintrón-Valentín ^a, Lorenzo García-Amaya^b and Nick C. Ellis^a

^aDepartment of Psychology, University of Michigan, Ann Arbor, MI, USA; ^bDepartment of Romance Languages and Literatures, University of Michigan, Ann Arbor, MI, USA

ABSTRACT

Previous research has demonstrated the effectiveness of captioning in second language (L2) comprehension and vocabulary learning. However, little attention has been paid to its potential in supporting grammar development, another challenging area in L2 acquisition. In this research, we used a randomised control design to investigate the role of salience-raising through textually enhanced captioned video on L2 vocabulary and grammar development. The experiment for this study was integrated into a one-semester university L2 Spanish grammar course. Through four data-collection sessions, we targeted: the preterite/imperfect, copula verbs, *gustar*-type verbs, and the subjunctive. In each session, participants saw a short grammar lesson before an animated video, in which participants were with; (i) captions that included textually enhanced target vocabulary, (ii) textually enhanced target grammar, or (ii) a no captioning control. Learners were then tested on their recognition and production of the target items. Results show clear effects of captioning on vocabulary performance. For some grammar structures, there were also positive effects of captioning on production, whereas for other grammar structures no such effect was uncovered. Altogether, these findings confirm the effectiveness of captioning on vocabulary, and illustrate specific difficulties related to grammar captioning, such as the role of structure-specific saliency in the grammatical input.

KEYWORDS


Captioning; textual enhancement; vocabulary acquisition; grammar acquisition; morphology; Spanish


Q2

1. Introduction

There is increasing interest in the use of captioned videos as a means to promote comprehension and vocabulary learning (Montero-Perez, Van den Noortgate, and Desmet 2013; Vanderplank 2010). Although captioned videos might have similar potential in supporting the learning of grammatical aspects of language, little attention has been paid to this area of learning in the Second Language Acquisition (SLA) research literature. Can captioned media be used to aid in the development of more complex linguistic forms deriving from grammar? This study investigates captioned video in combination with visual-input enhancement techniques as a means of facilitating the acquisition of various aspects of verb morphology in the second language (L2) Spanish classroom.

Technological advances now make it possible for the integration of multimedia learning materials such as videos, television programmes, and the internet in L2 classrooms (Blake 2013; Plass and Jones 2005; Vanderplank 2010). Such materials are intended for classroom learners who do not otherwise receive the amount of target input necessary to achieve high levels of proficiency in their L2 (Blake 2013). While L2 immersion experiences in a country where the target language is spoken can make

CONTACT Myrna Cintrón-Valentín  mcintron@umich.edu

 Supplemental data for this article can be accessed <http://dx.doi.org/10.1080/09571736.2019.1615978>.

© 2019 Association for Language Learning

up for this lack of sufficient input, not all learners have the time or the resources to engage in such experiences. One way forward, therefore, might be the inclusion of technological resources within the L2 classroom designed to provide learners with additional opportunities for target-language contact. Two multimedia video resources that have received attention within SLA research are captioning and
 55 subtitling. Captioning provides intralingual subtitles where the text is presented in the same language as the audio. On the other hand, subtitling involves the presentation of the L1 translation of L2 audio (Jung 1990). In this paper, we focus on the effect of captioning as it more closely resembles authentic target-language exposure, and because of its demonstrated benefits in vocabulary acquisition.

2. Literature review: captioning research, input enhancement, and second language acquisition

Captioning was first introduced to television programming around the 1980s with the original intent of making this type of media more accessible to the hearing-impaired. However, realising the potential of this resource for other target populations, educational researchers began investigating the benefits of captioning for developing L2 language skills in both hearing children and adults. The early research on captioning primarily focused on determining if captioned video was better than non-captioned video in (i) improving learner comprehension of the video content (e.g. Garza 1991; Markham 1989, 1993, 1999; Price 1983), and (ii) promoting vocabulary learning (e.g. Huang and Eskey 1999; Neuman and Koskinen 1992). These two areas remain the focus of current research (e.g. Muñoz 2017; Sydorenko 2010; Winke, Gass, and Sydorenko 2010; Winke, Sydorenko, and Gass 2013), in which a variety of comprehension and vocabulary measures are used. A recent meta-analysis of such studies by Montero-Perez et al. (2013) confirms significantly large effects of captioning on
 75 listening comprehension ($g = 0.99$) and on vocabulary learning ($g = 0.87$).

On the benefits of captions for L2 vocabulary learning, some researchers have suggested that the presentation of multimodal input (e.g. aural, written and visual) through same-language captioning 'may help the foreign/second language learner associate the aural and written forms of words more easily and quickly than video without subtitles' (Borras and Lafayette 1994: 70). Winke et al. (2010) attribute the usefulness of captioned media to matters of attention, suggesting that this medium can help draw learners' attentional focus to unknown word forms, and promote subsequent noticing and learning through repeated exposure. This hypothesis is consonant with foundational theories in SLA, which stress that attention is central to successful L2 acquisition (e.g. Gass, Spinner, and Behney 2017; Schmidt 2001; Tomlin and Vila 1994). Schmidt's (2001) Noticing Hypothesis, for instance, holds that conscious attention to linguistic forms in the input is an important precondition to learning – 'people learn about the things they attend to and do not learn much about the things they do not attend to' (Schmidt 2001: 30). Vanderplank's (2016) model of language acquisition through captioned media similarly emphasises how the 'taking out' of language from captioned videos – the first step in acquiring target-language output – promotes learners' attention to language and allows them to shift their attentional focus in order to meet their learning goals through a process of adaptation.

Captions might serve to make L2 features more salient in the input and thus increase their probability of being attended. The role of salience¹ as it relates to the perceptual distinctiveness of a linguistic cue in the input has received increasing interest in recent years (Ellis 2006, 2017; Gass et al. 2017; Wulff and Ellis 2018): 'salient items or features are attended, are more likely to be perceived, and are more likely to enter into subsequent cognitive processing and learning' (Ellis 2017: 21). Montero-Perez et al. (2014) examined the role of salience in the captioning line by comparing (i) the absence of captions, (ii) standard captioning with full captions, (iii) full captions plus highlighted keywords, and (iv) keyword-only captions, for their effects on comprehension and vocabulary learning in L1-Dutch intermediate learners of French. Their results revealed that type of captioning did not
 95 affect comprehension scores, but did significantly affect vocabulary learning, with keyword-only captions and full-captions-plus-highlighted-keywords having the greatest effect over the no-captions

control on some measures of vocabulary learning involving recognition of form and meaning (but not production). Thus, captions can make vocabulary more salient for learners and promote the learning of form-meaning connections.

105 Salience-raising through visual manipulations in the captioning line might likewise be relevant to the learning of L2 grammar. Despite the vast availability of grammatical forms in the input, L2 learners quite often ignore certain aspects of morphological structure and focus more on the meanings of open-class words, such as nouns, verbs, adjectives and adverbs (e.g. Bardovi-Harlig 1992; Clahsen and Felser 2006; Schmidt 2001). L2 grammar is particularly challenging for learners because morphological forms are less salient in the physical input while, at the same time, their functional interpretations are less clear than the one-to-one mappings typical for vocabulary (DeKeyser 2005, Ellis 2017; Goldschneider and DeKeyser 2001). The use of salience-inducing Input Enhancement manipulations (Doughty and Williams 1998; Sharwood-Smith 1993) to promote attention to low salience grammatical features in written input has been well documented in the SLA literature (Han, Park, and Combs 2008; Lee and Huang 2008; Leow and Martin 2017). Textual enhancement (Sharwood Smith 1993; henceforth TE), for instance, uses visual manipulations, such as color-coding, boldfacing and underlining, to enhance forms in the written input, and therefore facilitate learners' further processing of these cues. Crucially, Lee and Huang (2008) review studies of TE and conclude that there are conflicting findings regarding its effectiveness. They suggest that these discrepancies may be explained by factors as a learner's L1 and L2, learner proficiency, the type, complexity and communicative value of target forms, treatment intensity, and the measures used to assess noticing and processing of these forms.

110 In the grammar-learning literature, TE has generally been limited to unimodal mediums, that is, it focuses on the enhancement of grammatical cues through written mediums only, in the absence of pictorial or aural cues. One exception is a recent study by Lee and Révész (2018), which investigated the effects of TE on the learning of pronominal anaphoric reference in L1 Korean learners of English through a series of multimodal input-based activities. However, this study did not directly investigate captioned videos nor did the authors provide learners with pictures aimed at directly guiding the narrative presented through the bimodal input (aural and written). To our knowledge, little or no work has been done to assess if captioned video can be effective in aiding acquisition of L2 grammar, or more specifically, if there are differential effects based on the grammatical structures in question. This is one of the primary objectives of the current study.

135 3. The present study

The current study aimed to extend previous research on captioning and second language acquisition by targeting grammar. The study had three specific aims:

- 140 (1) to examine the effects of full captions + TE vocabulary on improving learner knowledge of target vocabulary.
- (2) to examine the effects of full captions + TE grammar on improving learner knowledge of target grammatical forms.
- (3) to investigate if any initial gains of full captions + TE grammar on the production of grammar are maintained over time.

145 We included RQ1 into our design (i.e. inclusion of a Vocabulary group) in order to ensure replicability of previous findings of captioning on vocabulary acquisition. In addition, we wanted to utilize any effects on vocabulary as a benchmark against which the efficacy of grammar captioning could be assessed. This was a critical component to our methodology, since this is one of the first studies that enters the under-explored research domain focusing on the effect of captioning on grammar development.

We investigated the effects of TE within the captioning line in three experimental conditions: a No-Captions Control group which received L2 audio but no material in the captioning line; a Captions + TE Vocabulary group, in which target vocabulary were made salient; and a Captions + TE Grammar group, designed to raise the learner's awareness and attention to grammatical cues. We targeted four grammatical topics: (1) preterite/imperfect forms, (2) *ser/estar* (i.e. copula verbs), (3) *gustar*-type verbs, and (4) the subjunctive in noun clauses.² Each video additionally included target vocabulary.

4. Method

4.1. Participants

A total of 176 English-speaking L2 learners of Spanish were recruited from a Spanish Grammar course at a large Midwestern University in the USA. They were fifth-semester intermediate learners of Spanish, and participated in the study for credit as part of one of their course requirements.³ The course contained 12 sections, which were quasi-randomly assigned to one of three conditions: a No Captions group (Control); a Captions + TE Vocabulary group; and a Captions + TE Grammar group (see Table 1 for descriptive statistics). Of these participants, 39 (Control = 14; Vocabulary = 11; Grammar = 14) were excluded from the study (1) if they had been exposed to the Spanish language before age 6 ($n = 26$); (2) if they had participated in a L2 Spanish study-abroad experience for two months or more ($n = 16$; 9 participants overlapped with those who had been exposed to Spanish from an early age); or (3) if they missed multiple lab sessions ($n = 8$; 2 participants overlapped with those above).

4.2. Written instruments

4.2.1. Language history questionnaire

Participants completed a Language History Questionnaire (Li et al. 2014), which included basic demographic questions, and more thorough questions about their experience with different languages.

4.2.2. Spanish vocabulary proficiency test

The Lextale-ESP (Izura, Cuetos, and Brysbaert 2014), a 90-item (60 words + 30 non-words) Spanish vocabulary proficiency test was administered to all participants. In this test, participants were asked to select words they recognised as Spanish words. As recommended by Lemhöfer and Broersma (2012) and Brysbaert (2013), the test was scored using the following formula:

$$\text{Score} = N_{\text{yes}' \text{ to words}} - 2 * N_{\text{yes}' \text{ to nonwords}}$$

This scoring formula penalised for guessing behaviour, so that a participant who marks all words and nonwords as known, or one who answers randomly, would receive a score of 0 (learners were informed of this scoring protocol prior to partaking in the task). The Cronbach's alpha of this test as reported in Izura et al. (2014) is $\alpha = 0.96$ ($N = 90$).

We additionally included the experiment's target vocabulary in this test in order to control for any possible familiarity with these words. The target vocabulary were coded and scored separately. Participants received one point for each target word they recognised as Spanish, for a total of 25 points.

Table 1. Descriptive statistics.

Group	N subjects	Age range		Mean age (SD)	Sex		
		Minimum	Maximum		Females	Males	Unspecified
Control	63	17	29	19.02 (1.6)	36	24	3
Vocabulary	59	17	28	18.69 (1.6)	36	23	–
Grammar	54	18	23	18.61 (0.9)	38	16	–

4.2.3. Spanish grammar proficiency test

A 45-item grammar proficiency test (García-Amaya 2012) was additionally administered to the participants. The test consisted of a short passage with a series of multiple-choice fill-in-the-blank options, which presented grammatical items varying in complexity. Participants received one point for each correct response for a total of 45 points. We evaluated the reliability of this test using Cronbach's alpha and found it to be acceptable ($\alpha = 0.73$; $N = 137$).

4.2.4. Immediate posttests

Vocabulary recognition test. Participants were tested on their recognition of target vocabulary (see Table 2). They were presented with a series of written words and were asked to select 'True' if they recalled being exposed to that word in the experimental session, or 'False' if they did not recall the word. All 25 target words were tested as well as an additional 25 foils. A score of 1 was given for each correctly identified target word. The Cronbach's alpha of the test was $\alpha = 0.83$ ($N = 125$).

*Vocabulary translation test.*⁴ A translation test required learners to provide the Spanish translation of specific English words. Each correct translation was given a score of 1, as were productions that were off by just one or two letters, for example, *alberco* when the correct form was *alberca* 'pool', or *frentos*, when the correct form was *frenos* 'braces'. Synonyms not presented in the video were scored as incorrect. The Cronbach's alpha of the test was $\alpha = 0.90$ ($N = 37$).

Grammar recognition test. Participants were tested on their recognition of target grammatical forms. They were presented with multiple sentences and were instructed to select the correct verb form out of two possible options. A score of 1 was given for each correct identification. The Cronbach's alpha of the test was $\alpha = 0.51$ ($N = 114$).

Grammar translation test. A translation test presented participants with sentences in English and asked them to type the appropriate Spanish translation. The responses were scored according to the provision of the correct target inflection. For instance, for lab session 1, which targeted the preterite/imperfect, participants needed to distinguish the usage of the two past forms. The Cronbach's alpha of the test was $\alpha = 0.81$ ($N = 83$).

Table 2. Vocabulary targets and frequency information.

Word	Session	Word type	NIM Frequency
Emparedado	1	Noun	0.18
Sombrilla	1	Noun	4.26
Alberca	1	Noun	1.07
Sandía	1	Noun	1.07
Sigiloso	1	Adjective	2.13
Lancha	1	Noun	1.95
Frenos	1	Noun	–
Repisa	2	Noun	2.31
Pashmina	2	Noun	–
Confites	2	Noun	0.36
Chucho	2	Noun	3.38
Impuntual	2	Adjective	0.18
Aulario	2	Noun	–
Dormilonas	3	Noun	–
Caniches	3	Noun	0.36
Sobremesa	3	Noun	0.71
Impúdico	3	Adjective	0.89
Espejuelos	3	Noun	–
Holgazán	3	Adjective	0.18
Estantería	4	Noun	2.66
Vergel	4	Noun	1.07
Alambrado	4	Noun	0.36
Boceto	4	Noun	1.07
Valija	4	Noun	2.31
Atolondrado	4	Adjective	0.18

Note. Session 1 = preterite and imperfect; session 2 = *ser* and *estar*; session 3 = *gustar*-type verbs; session 4 = subjunctive in noun clauses. Vocabulary words that do not include Frequency information are target words that were selected from a regional dialect.

4.2.5. Two-week delayed posttests

Approximately two weeks after each of the four experimental sessions, similar versions of the grammar translation tests were administered during learners' regular class session in order to measure retention over time. The tests included the same verb items the learners had been tested on in the immediate posttests, but in different sentence contexts. We included the grammar translation test only in the delayed posttest design, due to time constraints during the regular class sessions in which they were administered. The Cronbach's alpha of the test was $\alpha = 0.53$ ($N = 55$).

4.3. Grammar lesson videos

For each of the four target grammatical structures, a short video lesson was created. Each video lesson summarised how the relevant target form is conjugated in Spanish, provided learners with detailed discussions on two to three rules or verb instances, and included two to three practice exercises (See Figure 1). During each lab session, the video lessons were presented prior to the presentation of the animated videos.

4.4. Animated videos

Typically, in the captioning and vocabulary learning literature, the audiovisual materials consist of authentic video segments from diverse genres (e.g. documentaries, animated cartoons). In the current study, given our focus on specific grammar structures, we created our own animated videos. This included the process of generating original scripts for each target grammar structure, the recording of the characters' voices, and the animation of these scripts. This process allowed us to control for the frequency of occurrence of each of the vocabulary and grammar items, as well as their placement and randomisation in each of the videos.

The animated videos were created using Nawmal (www.nawmal.com), an animation programme that allows users to create videos by choosing from a menu of predesigned characters and sets. This software allows for much flexibility in the design, including the ability to upload user-recorded voices directly into the application, that are then automatically lip-synched to fictional characters. The software also allows for the the inclusion of hand and face gestures as the characters go through their dialogue, which can help make the scenes feel more dynamic and natural.

A total of four unique animated videos were created, one per target structure. For each structure, there were three versions of the video, which differed only in the focus of their captioning lines (Control, Captions + TE Vocabulary, or Captions + TE Grammar). For each video, captions were added using SRT Edit Pro (www.finalsub.com/sep.html), which allowed for the inclusion of color-coding and bold-facing within the captioning line.

As a measure of student engagement with the animated videos, learners responded to an exit survey, which consisted of open-ended questions about the usefulness of the animated videos, as well as questions regarding feedback for the improvement of the videos. A total of 143 out of the 176 learners (81%) who participated in the study completed the survey. On the whole, learners provided more positive (total = 126 (83%)) than negative (total = 26 (17%)) comments regarding the usefulness of the animated videos. Given the positive reception and engaged interest of the animated videos, we believe our materials to be adequate educational tools for learners at this level of instruction.

4.4.1. Vocabulary content

The animated videos created for each lab session included target vocabulary – overall a total of 25 target words were included in the experiment (see Table 2). The target vocabulary chosen for the experiment were either low-frequency words taken from the NIM Frequency database⁵ (Guasch et al. 2013), or regional vocabulary to which participants would have only been exposed if they

305

310

315

320

325

330

335

340

345

350

DISTRIBUCIÓN DE LOS VERBOS EN ESPAÑOL

INGLÉS

"TO LIKE" ES
UN VERBO
NORMAL

ESPAÑOL

LOS VERBOS
QUE FUNCIONAN
COMO "GUSTAR"
SON PARTICULARES

FUNCIONAMIENTO DE *GUSTAR* EN ESPAÑOL

ESTRUCTURA DE LA ORACIÓN

PREP.	+	OBJETO	+	PRONOMBRE OI	+	VERBO	+	SUJETO
A								
		Javier		le		gustan		las tostadas

PRÁCTICA 1

¿De las diferentes opciones qué estructuras son las correctas?



A Lourdes no me le gusta gustan el arte,

pero me le encanta encantan su perro.





Colour online, B/W in print

Figure 1. Illustrated above are representative slides taken from the ‘Gustar’- type verbs session. All lab sessions followed a similar structure. All participants, regardless of their experimental condition were first exposed to a short grammar lesson highlighting basic information on how each structure worked. Participants were additionally provided with two practice exercises.

were highly familiar with Puerto Rican or Mexican varieties of Spanish. This was done in order to control for learner familiarity of the target vocabulary. For each animated video, there were as many unique target words as there were grammar rules being targeted. For instance, for the preterite/imperfect session, there were seven vocabulary targets, the same number of grammar rules presented in the video. Each of the target words was presented four times, and though the unique items

were spread across the script, all repetitions of each word were massed (i.e. placed one after the other in consecutive sentences).

4.4.2. Grammar content

The specific grammar rules included in each video were taken from the course textbook *Repase y escribe: Curso avanzado de gramática y composición* (Cantelis Dominicus and Reynolds 2014). Depending on the target structure, either two or three rules, and one verb item representing each of the targeted rules, were first presented in the grammar lesson preceding the animated video. These same items, as well as the remaining rules and the verbs which exemplified them, appeared in the animated video. In the following sections, we provide an overview of the targeted structures (see Supplementary Material; Appendix A for full descriptions).

Session 1: Preterite and Imperfect. The standard usage of the Spanish past-tense system requires that learners understand the aspectual distinction between the preterite and imperfect (Comajoan 2013). Preterite forms characterise past actions as having a definitive beginning and endpoint (e.g. *caminé* 'I walked'), whereas imperfect forms characterise past actions or states being viewed as in progress (e.g. *caminaba* 'I was walking / I used to walk').

Each rule was represented through four different verb instances. Given that the acquisition of these structures in L2 Spanish can be influenced by lexical aspect (Bardovi-Harlig 2000), our design controlled for this variable in the selection of the preterite and imperfect verbs (see Supplementary Materials; Tables B1-B2, for the complete selection of rules and verbs, and Appendix C1 for an excerpt of the script).

Session 2: Ser and estar. Contrary to English, which only has one copula verb, 'to be', Spanish has two forms, *ser* and *estar*. The standard usage of these forms requires learners to understand the semantic properties that differentiate them. The current study targeted precisely the type of rules highlighted in the aforementioned studies, in other words, rules that we know learners are able to internalise during interlanguage development.⁶

For *ser/estar*, three rules for each form were included in the animated video. Each rule was represented four times, with *ser* and *estar* verbs conjugated in the first, second and third person singular (see Supplementary Materials; Tables B3-B4, for summary of rules and verbs, and Appendix C2 for an excerpt of the script).

Session 3: Gustar-type verbs. L1 learners' mastery of the *gustar*-type verb construction is considered especially challenging given the differences between its English counterpart 'to like' (e.g. Cerezo, Caras, and Leow 2016). Whereas 'like' codes as subject the entity that experiences a certain feeling, and as object the stimulus responsible for that feeling, *gustar* expresses the experienter though an indirect object (or dative) and the stimulus through the subject" (Vázquez Rosa 2006: 1). Although previous literature on the acquisition of *gustar*-type verbs focuses on the processing and use of the clitic pronoun preceding the verb (e.g. Lee and Malovrh 2009), in our study we focus specifically on an additional, sometimes overlooked challenge in the acquisition of these structures, namely the agreement between verb morphology and its subject.

We included six different verbs – *gustar* 'to like', *encantar* 'to love', *interesar* 'to be interested', *importar* 'to care', *molestar* 'to be bothered', and *quedar* 'to be left' – each presented four times, twice in the singular form, and twice in the plural form (see Supplementary Materials; Appendix C3 for an excerpt of the script).

Session 4: Subjunctive in noun clauses. The Spanish subjunctive mood is typically used in sentences with multiple clauses, in which the subject of the main clause exerts influence or will on the subject of the subordinate clause, in this case, a noun clause that serves as the object of the verb (Gudmestad 2012). The subjunctive in L2 Spanish is often described as a 'late-emerging item in both first and second language learners' given its low frequency, and the low perceptual salience of the subjunctive inflection in the input (DeKeyser and Prieto Botana 2013: 454; Collentine 2013). However, studies have shown that breaking down the syntactic and inflectional components of this structure can facilitate its acquisition regardless of learners' readiness (Collentine 2013). To this end, in the current study,

Table 3. Summary of the TE-captions manipulations per Grammar Topic.

Grammar topic	Condition		
	Control	Captions + TE Vocabulary	Captions + TE Grammar
Preterite/imperfect	n/a	Target vocabulary is bold	Target verb is bold and yellow
<i>Ser/estar</i>	n/a	and yellow	Target verb is bold and yellow
<i>Gustar</i> -type verbs	n/a		Target verb is bold and yellow
Subjunctive in noun clauses	n/a		Main clause verb, conjunction <i>que</i> 'that', and subordinate subjunctive verb are in bold and yellow

both the verb in the main clause, which acts as a cue to the subjunctive, and the subordinated subjunctive verb, were made salient in order to facilitate learners' understanding of the rules underlying subjunctive usage. For the subjunctive in noun clauses, five rules were targeted. Each rule was represented by four different verb instances (see Supplementary Materials; Tables B5-B6 for summary of rules and verbs; and Appendix C4 or an excerpt of the script). Twelve indicative sentences were included as fillers (see Supplementary Materials; Table B7).

4.4.3. Captioning content and textual enhancement manipulations

The effect of TE on vocabulary and grammar within the captioning line was investigated through three experimental conditions: Control, Captions + TE Vocabulary and Captions + TE Grammar (see Table 3 for a summary of these conditions by structure) Figure 2.

4.5. Data collection procedure

On the first day of class of the 15-week semester, two members from the research team attended all 12 course sections and administered the two Spanish proficiency tests.⁷ The experimental phase of the study took place over four different sessions spaced through the semester in the order presented in the course syllabus: (1) preterite/ imperfect, (2) *ser/estar*, (3) *gustar*-type verbs, and the (4) subjunctive in noun clauses. Due to the curricular constraints of the grammar course, it was not possible to counterbalance the presentation of the four grammar topics. During each 50-minute session, the experimenters met with the learners on their assigned class day and time in a pre-assigned computer classroom.⁸ The experimental protocol was computerised and made available to each participant through the Canvas Learning Platform (<https://www.canvaslms.com/>). During each experimental session, learners were presented with the grammar lesson video, followed by the corresponding animated video manipulated for one of three conditions. Following the two videos, participants completed the vocabulary recognition and production tests outlined in section 4.2.4. Delayed posttests on each grammar structure were administered approximately two weeks following their corresponding session (see Table 4 for an overview of the procedure).

4.6. Statistical analysis

Statistical analyses were conducted using R Studio version 1.0.143 (RStudio Team 2015). The data were analysed through generalised linear models and multilevel generalised linear regression models utilising the `glm()` and `glmer()` functions within the `lme4` package in R (Bates, Maechler, and Bolker 2015).

4.6.1. Vocabulary data

For the vocabulary recognition and translation analyses, we ran logistic regression models on the pooled results (collapsing across all vocabulary sessions). The dependent measure was proportion of trials correct, with GROUP (Control, Captions + TE Vocabulary and Captions + TE Grammar) as the predictor term. The week 1 VOCABULARY PROFICIENCY test was additionally included as a fixed variable

455

460

465

470

475

480

485

490

495

500

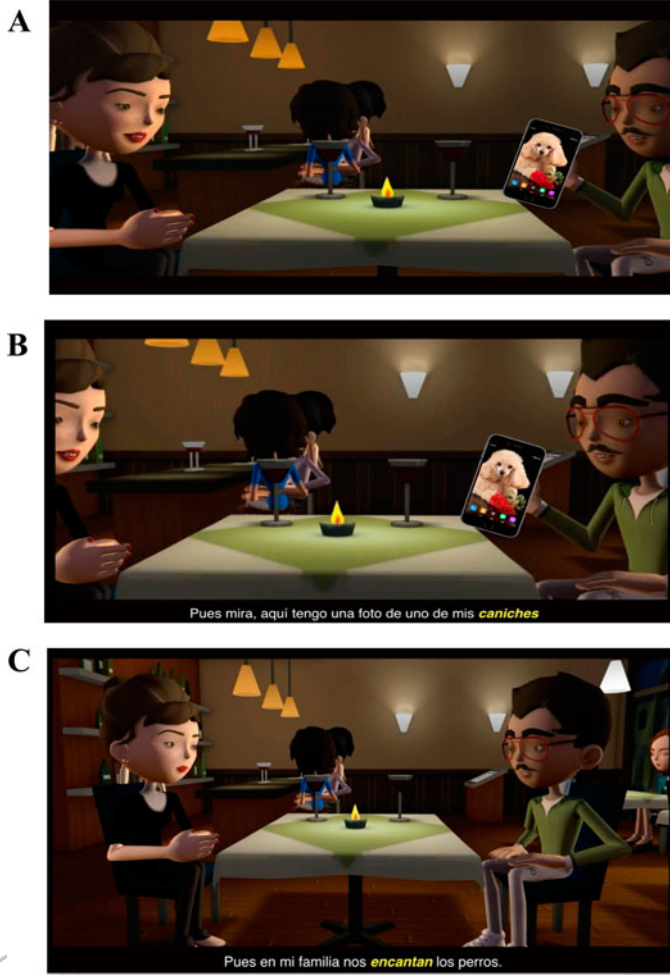


Figure 2. Illustration of all three condition types taken from the *gustar*-type verbs session. (A) Participants in the Control group did not receive subtitles in their treatment. (B) Participants in the Vocabulary group received subtitles where the target vocabulary was highlighted in bold and yellow. (C) Participants in the Grammar group received subtitles in which the target grammatical structures were highlighted in bold and yellow.

Table 4. General overview of procedure.

Phase	Test	Time
Pre-experimental Phase	Grammar Proficiency Test	First day of class
	Vocabulary Proficiency Test	First day of class
	Language History Questionnaire	First week of class
Experimental Phase	Grammar Video Lesson	Experimental session (4 times)
	Animated Video	
	Immediate Vocabulary Recognition	
	Immediate Vocabulary Translation	
	Immediate Grammar Recognition	
	Immediate Grammar Translation	
Two-week delayed Grammar Translation Test	Two-week in-class posttest (4 times)	

Note. The Experimental Phase took place during eight different time points across the 15-week semester. Students saw the animated videos and took the immediate posttests for each of the four structures on their assigned class day. Two-weeks after each experimental session, participants were tested on their production of the grammar structure.

Colour online. B/W in print.

to take into account individual differences in Spanish proficiency. This variable was mean-centered before being added to the model.

4.6.2. Grammar data

For the grammar recognition and translation analyses, we fit logistic regression models to the repeated count measures using the `glmer()` function. The dependent variable was the binomial count of correct trials, offset by the total number of trials for each respondent (given our objective of modelling the probability of a correct trial), with `GROUP`, `STRUCTURE` (preterite/imperfect, *ser/estar*, *gustar*-type verbs and subjunctive), and their two-way interaction as predictors. We accounted for the expected correlation of the repeated counts for each subject by including random subject effects in the models. The week 1 `GRAMMAR PROFICIENCY` measure was additionally included as a fixed variable to take into account individual differences in Spanish proficiency. The `GRAMMAR PROFICIENCY` measure was mean-centered before being included in the model.

We used a maximal random effects modelling procedure, following the advice of Barr et al. (2013). This modelling included by-subjects random intercepts and by-subjects random slopes for the predictor `STRUCTURE`. In order to decide between converging models, we retained the most complex model with the lowest AIC and BIC terms that converged after 10,000 iterations under this procedure.

4.6.3. Missing data

For three sections, participants' data for the preterite/ imperfect session was treated as missing data because they saw the animated video more than once (Control = 15; Captions + TE Vocabulary = 16; Captions + TE Grammar = 17). This was also the case for participants who were absent from any of the four lab sessions and who were presented with the lab material by their instructor before their make-up session ($n = 17$).

For each participant, any experimental word known at baseline was treated as missing for the vocabulary recognition data. This was not done for the vocabulary translation data given that the initial baseline measure of recognition is not an accurate reflection of the participants' ability to translate these words. This information was extracted from the initial Spanish vocabulary proficiency test where we included all of the experimental words as a baseline measure of their knowledge of these forms (see section 4.2.2).

5. Results

5.1. Proficiency data

Table 5 presents the group means, standard deviations and confidence intervals for the Vocabulary and Grammar proficiency tests administered on the first day of class.

The vocabulary proficiency test included 25 items that were used as the target vocabulary items in this study. These items were removed from the scoring of the proficiency test to separately assess learners' prior knowledge of these words.

Table 5. Means, standard deviations (SD), and 95% Confidence Intervals (CI) for the Vocabulary and Grammar Proficiency Tests.

Group	Mean	SD	95% CI
<i>Group Accuracy scores for Vocabulary Proficiency</i>			
Control	6.489	7.843	[4.294, 8.686]
Vocabulary	9.000	8.145	[6.696, 11.304]
Grammar	8.900	5.939	[7.059, 10.741]
<i>Group Accuracy scores for Grammar Proficiency</i>			
Control	24.367	4.915	[22.991, 25.743]
Vocabulary	25.043	5.213	[23.568, 26.517]
Grammar	23.550	4.771	[22.071, 25.029]

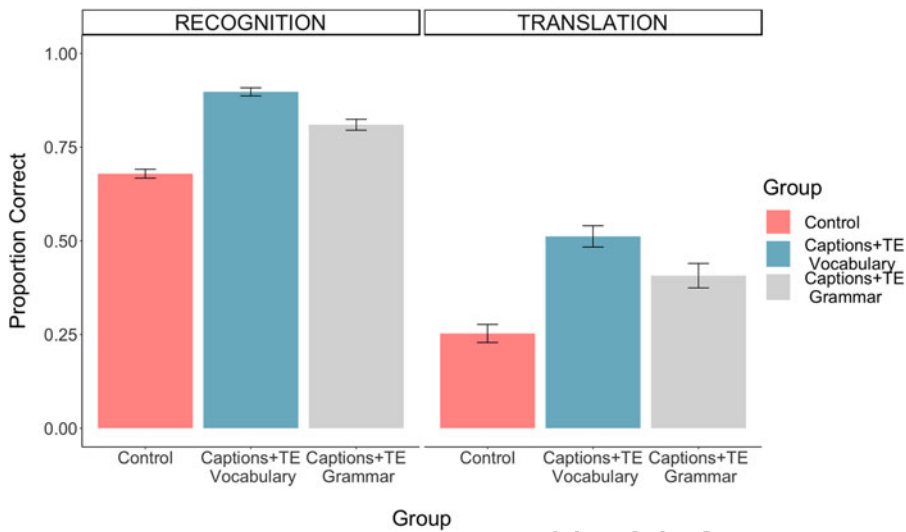


Figure 3. Mean Accuracy Scores for Vocabulary Recognition (left panel), and Vocabulary Translation (right panel). Error bars are 2 standard errors long.

5.2. Vocabulary recognition

The Vocabulary recognition post-test data are plotted in the left-hand panel of Figure 3 (see Supplementary Materials; Table D1). The pattern for the recognition data suggests an advantage of captioning over non-captioned video, with both captioning groups scoring higher than the no captions Control group. Additionally, the data patterns suggest an overall advantage for the Captions + TE Vocabulary group participants over the Control and Captions + TE Grammar groups (see the left-hand panel of Figure 3 and the top panel of Table 6 for details). To investigate the effects of captioning, we ran a generalised linear model which included fixed effects of VOCABULARY PROFICIENCY and our main variable of interest: GROUP. The first model, with the Control group as the reference level, revealed significant positive group effects, when comparing to both the Captions + TE Vocabulary, $\beta = 1.352$, $SE = 0.087$, $p < 0.001$ and Captions + TE Grammar groups, $\beta = 0.658$, $SE = 0.078$, $p < 0.001$. Thus, both captioned groups were more accurate in their recognition accuracy than the controls. The same model, with Captions + TE Vocabulary as the reference level, revealed a significant negative group effect, when compared to the Control $\beta = -1.352$, $SE = 0.087$, $p < 0.001$ and the Captions + TE Grammar groups, $\beta = -0.694$, $SE = 0.094$, $p < 0.001$. Thus, there was an advantage of the Captions + TE Vocabulary group over the Captions + TE Grammar group and the Control Group in their recognition accuracy.

Table 6. Vocabulary Recognition Result summary: coefficient estimates β , standard errors $SE(\beta)$, associated Wald's z-score ($=\beta/SE(\beta)$) and significance level p for all predictors in the analysis.

Predictor	Coef. β	$SE(\beta)$	z	p
Model 1 with the Control group as the reference level				
(Intercept)	0.774	0.0491	15.770	<0.001***
Vocabulary Group	1.352	0.087	15.588	<0.001***
Grammar Group	0.658	0.078	8.437	<0.001***
Vocabulary Proficiency	0.158	0.035	4.534	<0.001***
Model 2 with the Vocabulary group as the reference level				
(Intercept)	2.125	0.072	29.697	<0.001***
Control Group	-1.352	0.087	-15.588	<0.001***
Grammar Group	-0.694	0.094	-7.399	<0.001***
Vocabulary Proficiency	0.158	0.035	4.534	<0.001***

Table 7. Vocabulary Translation Result summary: coefficient estimates β , standard errors SE (β), associated Wald's z-score ($= \beta/SE$) and significance level p for all predictors in the analysis.

Predictor	Coef. B	SE (β)	z	p
Model 1 with the Control group as the reference level				
(Intercept)	-1.019	0.075	-13.545	<0.001***
Vocabulary Group	1.034	0.098	10.558	<0.001***
Grammar Group	0.524	0.105	4.989	<0.001***
Vocabulary Proficiency	0.317	0.040	7.803	<0.001***
Model 2 with the Vocabulary group as the reference level				
(Intercept)	0.014	0.063	0.224	n.s
Control Group	-1.034	0.098	-10.558	<0.001***
Grammar Group	-0.510	0.096	-5.316	<0.001***
Vocabulary Proficiency	0.317	0.041	7.803	<0.001***

5.3. Vocabulary translation

As in the vocabulary recognition results, the data pattern for the translation scores suggests an advantage of captioning over non-captioned video, as well as an overall advantage for the Vocabulary group over the Control and Grammar groups (see the right-hand panel of Figure 3, the bottom panel of Table 7 and Supplementary Materials; Table D1 for details). We ran the same analysis design as for the recognition data. The first model, with the Control group as the reference level, revealed a significant positive group effect, compared to Captions + TE Vocabulary, $\beta = 1.034$, $SE = 0.098$, $p < 0.001$, and to Captions + TE Grammar, $\beta = 0.524$, $SE = 0.105$, $p < 0.001$, i.e. both captioned groups were more accurate in their production accuracy. The same model, with Captions + TE Vocabulary as the reference level, revealed a significant negative group effect, compared to Control $\beta = -1.034$, $SE = 0.098$, $p < 0.001$, and to Captions + TE Grammar, $\beta = -0.510$, $SE = 0.096$, $p < 0.001$, confirming our initial observation of the overall advantage of the Captions + TE Vocabulary group.

5.4. Grammar recognition

Figure 4 illustrates the group mean scores as well as the standard errors for all four target grammar structures (see Supplementary Materials; Table D2). Here, the overall pattern does not suggest any clear group differences within each structure. We ran a generalised linear mixed effects model, which included proportion of trials correct as the dependent measure. Each model additionally included three fixed effects, two of which were predictor variables: STRUCTURE, and GROUP; and one of which was a control variable: GRAMMAR PROFICIENCY. We retained the most complex model with the lowest AIC and BIC terms that converged after 10,000 iterations under this procedure. In our data, this meant that we first tested all three-way combinations of levels within STRUCTURE. However, none of these models converged. Following this, we tested every possible two-way combination of levels within STRUCTURE – all of these models converged. We decided on a final model (from the latter set of converging models) by selecting the model that generated the lowest AIC and BIC terms (see Supplementary Materials; Table D3 for the full summary of the final model). Given that our design focused on whether there were differences between each captioning condition within each grammar topic, we ran multiple iterations of the same model using different reference levels for GROUP and STRUCTURE. Our initial observations were confirmed by our models, which did not reveal any significant GROUP by STRUCTURE interactions.

5.5. Grammar translation

5.5.1. Immediate posttest

The left-hand panels of Figure 5 plot the group mean scores as well as the standard errors by structure for the Immediate Posttests, showing varying effects of captioning on production by structure (see Supplementary Materials; Table D4). To investigate group differences, we ran a generalised

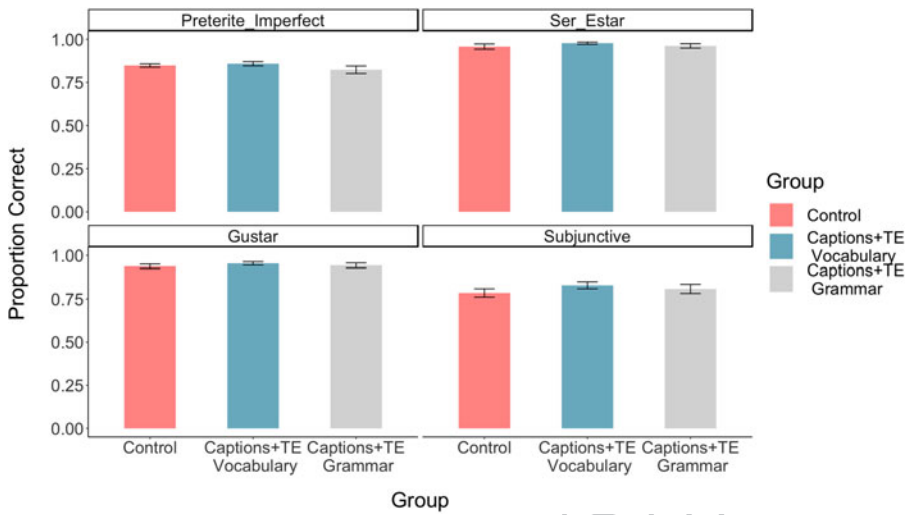


Figure 4. Mean Accuracy Scores for Grammar Recognition by Structure and Group. Error bars are 2 standard errors long.

linear mixed effects model which included fixed effects of GRAMMAR PROFICIENCY, and our main variables of interest: GROUP, and STRUCTURE as predictor terms. We followed the same analysis procedure outlined in section 5.4. The initial model with the lowest AIC and BIC terms revealed significant group by structure interactions for gustar-type verbs and the subjunctive (see Supplementary Materials; Table D5 for the full summary). Again, given that our design focused on whether there were differences between each captioning condition within each grammar topic, we ran multiple iterations of the same model using different reference levels for group and structure. When comparing the Control group against the Captions + TE Grammar group, there were significant differences for *gustar*-type verbs ($\beta = 0.496$, $SE = 0.203$, $p < 0.05$) and the subjunctive structures ($\beta = 0.503$, $SE = 0.206$, $p < 0.05$). The captioning in the Captions + TE Vocabulary group also had an effect on the subjunctive ($\beta = 0.525$, $SE = 0.199$, $p < 0.01$).

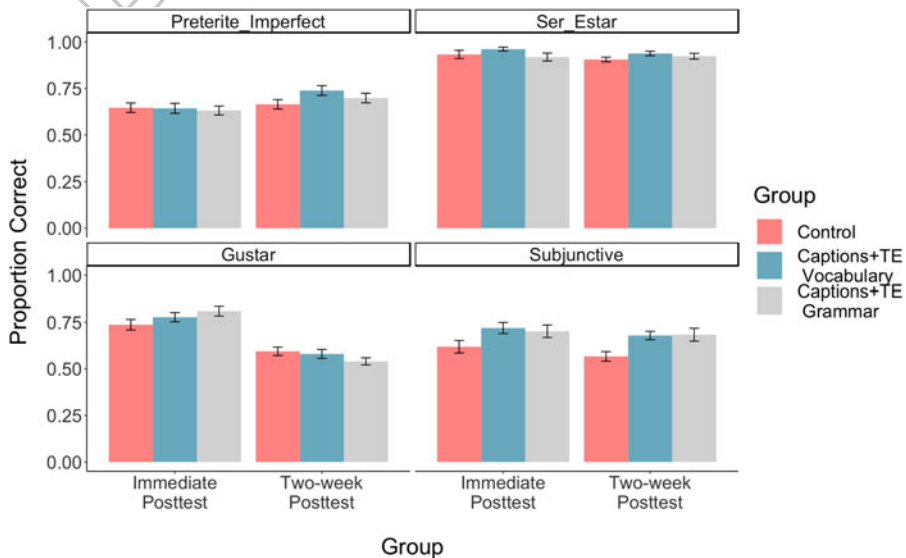


Figure 5. Mean Accuracy Scores for Grammar Translation by Structure, Group and Time. Error bars are 2 standard errors long.

5.5.2. Two-week posttest

The right-hand panels of Figure 5 plot the group mean scores as well as the standard errors by structure for the Two-week Posttests (see Supplementary Materials; Table D4). Again, the data pattern shows varying effects of captioning on production by structure. To investigate group differences, we followed the same procedure as for the Immediate Posttest data. The initial model with the lowest AIC and BIC terms revealed significant group by structure interactions for *gustar*-type verbs only (Supplementary Materials; Table D6 for the full summary). Our follow-up models revealed significant effects when comparing the Control group against the Captions + TE Grammar group for both *gustar*-type verbs ($\beta = 0.508$, $SE = 0.227$, $p < 0.05$) and the subjunctive structures ($\beta = 0.507$, $SE = 0.190$, $p < 0.01$). The captioning in the Captions + TE Vocabulary group also had an effect on the subjunctive ($\beta = 0.528$, $SE = 0.184$, $p < 0.01$).

To summarise the grammar results, the immediate posttest data show significant effects of captions on *gustar*-type verbs and on the subjunctive. The results for the two-week posttest reveal that these effects were maintained for both *gustar*-type verbs and the subjunctive.

6. Discussion

6.1. Vocabulary

The first aim of this study was to examine the effects of full captions + TE vocabulary on improving learner knowledge of target vocabulary. Our results showed positive effects of both captioning and of specific highlighting with TE. Specifically, the vocabulary recognition and production results show that learners in both captioning groups were more successful than non-captioned control learners in acquiring the target words. There was an effect of vocabulary TE on both the recognition and production scores. This is evidenced by the advantage of the Captions + TE Vocabulary group over both the Control and Captions + TE Grammar groups, which did not include highlighting on vocabulary.

These findings lend support to previous research demonstrating the role of captioning in promoting learner knowledge of L2 vocabulary (e.g. Montero-Perez 2013). It seems, that – at least for vocabulary – the provision of on-screen text with or without enhancement of the target vocabulary (in other words the Captions + TE Vocabulary and the Captions + TE Grammar groups), facilitates learners' 'taking out' of language (Vanderplank 2016). This may be due, in part, to the adaptation process described in Vanderplank (2016), whereby learners select 'the language attended to for their own purposes' (p. 239). In this case, the unfamiliarity of the target vocabulary could have led the learners in the Captions + TE Grammar group to isolate these lexical items as well. This is consonant with the notion of surprisal salience, where it is the infrequency of a particular word form that may lead to its increased prominence in the input (e.g. Gass et al. 2017). As part of our design, the target vocabulary selected for each of the videos were low in their frequency of usage (see section 4.4.1). In order to facilitate learner attention to these forms, we additionally manipulated the frequency of occurrence of the vocabulary within their corresponding videos. These two factors could have increased their salience in the input regardless of the focus of the TE manipulations. Specifically, upon first encounter of a given vocabulary item, learners' attention could have been drawn to the unknown word form given its infrequency, whereas the subsequent occurrences of the vocabulary word form could have allowed learners to gather further information about its meaning. In other words, our results suggest that, by visually enhancing target words in the captioning line, learners may be more able to isolate unknown word forms from the captioning line and make initial form-meaning connections.

The current findings partially confirm those of Montero-Perez et al. (2014). In Montero-Perez et al. (2014), learners in the salience conditions outperformed those in the control condition in a recognition task, but not in a production task. Contrastingly, in our study, learners in the salience conditions outperformed those in the control condition in both tasks. One possible explanation for the difference between studies may derive from the type of tests used in the two designs: our production

task required learners to translate the target words from their L1 to their L2, whereas in Montero-Perez et al. (2014), learners translated the target words from their L2 to their L1, a skill in which learners are typically faster and more accurate (e.g. Kroll and Stewart 1994). L1-to-L2 translation is a more discriminating task, one where multiple modalities of representation can usefully support the retrieval and production of the L2 form.

6.2. Grammar

Our second research aim was to examine the effects of full captions + TE grammar on improving learner knowledge of target grammar. We were additionally interested in investigating if any initial gains of full captions + TE grammar are maintained over time in grammar production (this was our third research aim). Contrary to the clear set of vocabulary findings, the findings for grammar were mixed. On the one hand, for recognition, no significant differences were found between the groups for any of the structures. On the other hand, for production, captioned videos showed an advantage over non-captioned videos for *gustar*-type verbs and the subjunctive in noun clauses (but not for *ser/estar* or the preterite/imperfect) – and this was true both for the immediate posttest and the two-week posttest. On the whole, the findings from the production task suggest that learner knowledge of some (but not all) structures is more easily enhanced by captioning.

The four grammar structures included in this study share one commonality in that they all required learners to supply appropriate verb conjugations. Of importance, however, is that each of these structures differs in terms of their specific semantic and syntactic properties and thus on the specific features that learners were expected to process from the input (see also section 4.4.2 and Supplementary Material; Appendix A). It is these structure-specific characteristics that may have additionally modulated their saliency and in turn learners' ability to extract particular patterns regarding their correct usage. In sections 6.2.1 through 6.2.4, we focus on the effects uncovered for each structure and provide a discussion on how structure-specific characteristics may have modulated their saliency in the input.

6.2.1. Preterite and imperfect

Regarding the preterite/imperfect forms, we did not uncover significant differences between the three experimental groups. As mentioned in section 4.4.2, the appropriate usage of these forms requires that learners understand how to encode past aspectual distinctions morphologically. This process involves internalizing a set of rules that describe the contexts in which each form is used. One possible explanation for present findings could therefore relate to the number of structures and rules included in a single session. Along these lines, Overstreet (1998) suggests that the lack of a TE effect on the acquisition of preterite/imperfect may be due to the difficulty of learning how two forms contrast within a specific semantic context (see also Jourdenais et al. 1995; Leeman et al. 1995). Overstreet suggests further that TE may be more effective when directed at one grammatical form at a time instead of the contrast between the two. Elaborating further on this idea, Han et al. (2008) note that although TE has been found to promote noticing and learning of some linguistic constructions, more research is needed to uncover whether these effects create an additional trade-off with comprehension both at the local and global levels. At the local level, TE on the preterite/imperfect forms might actually distract learners' attention from the surrounding discourse, which offers critical information regarding how the two aspectual choices are used (see also Bardovi-Harlig 1998, regarding the importance of narrative context). It is thus possible that in the context of our study, the number of rules being targeted, as well as the added TE on the forms, could have negatively modulated learners' attention away from the rules being presented.

6.2.2. Ser and estar

For *ser/estar*, learners were at ceiling, and significant differences were not uncovered between the experimental groups. Previous research on the *ser/estar* distinction has shown that the target-like

usage of these verbs is characterised by distinct developmental stages whereby the proper distinction of the two forms in different contexts is more prevalent at more advanced stages of learning (e.g. Van-Patten 1987). The learners in the current study were intermediate learners of Spanish; it is possible that they already had ample experience with the copula contrast in their L2 (see section 6.2 for a discussion of this limitation). Although we did not assess learners' prior knowledge in the present experimental design (e.g. through a pre-test), Lee and Huang (2008) suggest that TE might not make significant contributions to the learning of structures that are well-ingrained in learners' prior knowledge.

6.2.3. *Gustar-type verbs*

For *gustar*-type verbs, the results of the current study suggest that learner knowledge of subject-verb agreement can be supported by multimodal captioned media. As mentioned previously, appropriate subject-verb agreement in the context of this structure requires learners to understand the non-canonical mapping of thematic roles. Learners must additionally learn the set of verbs that require this type of construction (see section 4.4.2). Once acquired, learners need only apply the same rule to each verb instance. One possible explanation for the results uncovered here is that learners might have used the same type of learning strategies as they did for the learning of the vocabulary target words, hence the similar gains. In other words, following the grammar lesson, even learners who had little prior knowledge of this form should have had some understanding of how the structure works, realising that there are specific verbs in the Spanish language that follow the *gustar*-type pattern. It is possible that learners were primarily focused on identifying these specific *gustar*-type verbs rather than determining how to conjugate them within the sentential context.

An additional consideration relates to the nature of the experimental items included in the study. Specifically, Cerezo et al. (2016) categorise *gustar*-type structures on four levels (i.e. Types 1 through 4) according to their processing difficulty, which they define as 'the number of steps that L1 English speakers need to take to process or produce them' (p. 273). In Cerezo et al.'s framework, Type 1 *gustar*-type structures are considered the least complex and consist of structures where the experiencer is a first-person noun or second-person singular pronoun. In the current study, the majority of the experimental sentences presented in the animated videos and in the assessment tasks fell under the Type 1 category. Future studies could well apply the type of framework described by Cerezo et al. (2016) in order to more thoroughly investigate the scope of TE + captions in facilitating the acquisition of *gustar*-type structures.

6.2.4. *Subjunctive in noun clauses*

Similar to *gustar*-type verbs, the results for the subjunctive in noun clauses suggest that learner knowledge of L2 grammar can be supported by multimodal captioned media. As mentioned in section 4.4.2, the Spanish subjunctive is a relatively complex morpho-syntactic structure emerging late in both L1 and L2 Spanish acquisition. Nonetheless, studies have shown that breaking down the syntactic dependencies and inflectional components of this structure can facilitate its acquisition regardless of learners' readiness (Collentine 2013). In the current study, both the verb in the main clause, which can act as a cue to the subjunctive, and the subordinated subjunctive verb, were made salient to the learners. Although the competing effect of highlighting the main clause or the subordinated verb cannot be assessed given our research design, we would argue that it was the highlighting of the main-clause verb that more strongly facilitated learners' understanding of the subjunctive rules, as indicated in Farley and McCollam (2004) and Collentine (2013).

To summarise section 6.2, we have suggested that, within the domain of grammar learning in captioning and TE studies, it is essential to account for structure-specific characteristics that might modulate their saliency in the input. For the four structures examined here, we have brought to light the importance of taking into account factors such as: learners' prior knowledge of the structures in question (e.g. *ser/estar*); the discourse context (e.g. the preterite/imperfect); the number of rules being thought of during input processing (e.g. *gustar*-type verbs versus the preterite/imperfect); and how highlighting syntactic dependencies or additional contextual cues may facilitate the learning process (e.g. the subjunctive).

6.3. Limitations and future directions

One limitation of our study is that we did not counterbalance the order of grammar structures; this was not possible in our design since we followed a common syllabus for all sections of the grammar course. In future research, it will be beneficial to counterbalance the order of grammar structures to determine whether increased positive effects over time may be a product of participants' growing familiarity with the experimental procedure.

A second limitation is that we did not include a pretest prior to conducting the experimental sessions. Without a pretest, it is difficult to tease apart any possible confound regarding the gains acquired through the treatment from pre-existing knowledge. In our study, this issue becomes relevant when trying to interpret the near-ceiling effects for *ser/estar*. One question that we cannot answer, for example, is whether the learners already knew these structures well enough going into the study, or if the grammar lesson/animated video were sufficiently effective. Future work would benefit from the inclusion of a pretest/posttest design in order to discern the effects of prior knowledge from the experimental treatment.

Moving forward, our results underscore the importance of tailoring TE to specific target structures so that the appropriate inflectional, syntactic, and functional considerations are emphasized. It would thus be useful for future studies to assess effects of different designs of TE-captions on the structures in question. Future research should likewise consider the inclusion of research tools designed to measure learners' immediate noticing of perceptually enhanced input in addition to more traditional acquisition measures (e.g. Han et al. 2008: 601). The inclusion of methodologies such as eye-tracking (see for instance Lee and Révész 2018; Montero Perez, Peters, and Desmet 2015; Muñoz 2017) would allow for a more complete understanding of the potential interaction of salience, learner attention and TE-captioned media in L2 grammar development.

7. Conclusion

We have presented an innovative study aimed at exploring the role of TE + captioned media on the learning of vocabulary and grammar in the L2 Spanish classroom. We targeted four grammar topics: preterite/imperfect, *ser/estar*, *gustar*-type verbs, and the subjunctive. The findings for vocabulary showed positive effects of both captioning and of specific highlighting of target vocabulary with TE, with an overall advantage for the latter. For some grammar structures – namely, *gustar*-type verbs and the subjunctive – there were positive effects of captioning on production, whereas for other grammar structures no such effect was uncovered.

An important implication of the present findings is that L2 researchers and instructors should consider structure-specific characteristics that mediate the saliency of individual grammatical items when implementing TE-captioned media. Such considerations include: (i) learners' prior knowledge of the structures in question; (ii) the number of grammar rules involved in the learning process; (iii) the role of the discourse context in learners' processing of the target grammar; and (iv) how the highlighting of syntactic dependencies or additional contextual cues may facilitate the learning process.

Notes

1. We adopt this definition of salience due to the focus of the current study. We acknowledge that the topic of salience within the SLA literature is broad in scope. For instance, Cintrón-Valentín and Ellis (2016) focus on the physical characteristics of the linguistic cues in the input, learners' prior L1 knowledge, and Form-Focused Instruction techniques aimed at refocusing learner attention (see also Ellis 2017; Gass et al. 2017).
2. We focused on these grammatical structures since these were the four major grammar topics covered in the course, for which more than one day of class instruction was assigned. For all other grammar structures covered in the course (e.g., *por/para* 'for/to'), only one half-day of grammar instruction was included in the syllabus.
3. Participants were fifth semester learners of Spanish or had received a high score in their Advanced Placement Spanish course in high school.

4. The vocabulary translation task, as we call it here, has typically been referred to in the vocabulary learning literature as a test of form recall by Nation (2001) and as a productive translation task by Webb (2008).
5. NIM is Web-based software that allows users to search for words according to their length, lexical frequency, or parts of speech in English, Spanish, and Catalan.
6. Although we recognize that the usage of this structure is in variation and that this variation can affect its acquisition (Geeslin 2003), in the current study we focused on the rules included in the learners' course textbook.
7. Learners who were absent on the first day of class, or enrolled after the first week, completed the Pre-Experimental phase during a separate makeup session.
8. Instructors were asked not to assign readings or homework on the target material prior to the experimental phases.

Acknowledgement

We would like to thank Nicholas Henriksen and Elizabeth Garno for helping design and pilot parts of this experiment, as well as the Undergraduate Research Opportunities Program (UROP) at the University of Michigan.

Disclosure statement

Q6 No potential conflict of interest was reported by the authors.

Funding

This research was supported by National Science Foundation Graduate Research Fellowship under [grant number DGE 1256260]. We would like to thank Nicholas Henriksen and Elizabeth Garno for helping design and pilot parts of this experiment, as well as the Undergraduate Research Opportunities Program (UROP) at the University of Michigan.

ORCID

Myrna Cintrón-Valentín  <http://orcid.org/0000-0002-7560-7566>

References

- Bates, D., M. Maechler, B. Bolker and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67, no. 1: 1–48.
- Bardovi-Harlig, K. 1992. The use of adverbials and natural order in the development of temporal expression. *IRAL—International Review of Applied Linguistics in Language Teaching* 30, no. 4: 299–320.
- Bardovi-Harlig, K. 1998. Narrative structure and lexical aspect: conspiring factors in second language acquisition of tense-aspect morphology. *Studies in Second Language Acquisition* 20, no. 4: 471–508.
- Bardovi-Harlig, K. 2000. *Tense and Aspect in Second Language Acquisition: Form, Meaning, and use*. Oxford: Blackwell.
- Barr, D.J., R. Levy, C. Scheepers and H.J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language* 68, no. 3: 255–78.
- Q7** Blake, R.J. 2013. *Brave New Digital Classroom: Technology and Foreign Language Learning*. Georgetown University Press.
- Borrás, I. and R.C. Lafayette. 1994. Effects of multimedia courseware subtitling on the speaking performance of college students of French. *The Modern Language Journal* 78, no. 1: 61–75.
- Brysbaert, M. 2013. Lextale_FR a fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica* 53, no. 1: 23–37.
- Canteli Dominicis, M. and J.J. Reynolds. 2014. *Repase y Escribe: Curso Avanzado de Gramática y Composición*. 7th ed. New York: St. John's University.
- Cerezo, L., A. Caras and R.P. Leow. 2016. Effectiveness of guided induction versus deductive instruction on the development of complex Spanish 'gustar' structures: an analysis of learning outcomes and processes. *Studies in Second Language Acquisition* 38, no. 2: 265–91.
- Cintrón-Valentín, M.C. and N.C. Ellis. 2016. Saliency in second language acquisition: physical form, learner attention, and instructional focus. *Frontiers in Psychology* 7: 1284.
- Clahsen, H. and C. Felser. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27, no. 1: 3–42.
- Collentine, J. 2013. Subjunctive in second language Spanish. In *The Handbook of Spanish Second Language Acquisition*, ed. K.L. Geeslin, 270–88. Oxford: Wiley Blackwell.
- Comajoan, L. 2013. Defining and coding data: narrative discourse grounding in L2 studies. In *Research Design and Methodology in Studies on L2 Tense and Aspect*, ed. R. Salaberry and L. Comajoan, 309–56. Berlin: De Gruyter.

- DeKeyser, R.M. 2005. What makes learning second language grammar difficult? A review of issues. *Language Learning* 55, no. 1: 1–25.
- DeKeyser, R. and G. Prieto Botana. 2013. Acquisition of grammar by instructed learners. In *The Handbook of Spanish Second Language Acquisition*, ed. K.L. Geeslin, 449–65. Oxford: Wiley Blackwell.
- Doughty, C. and J. Williams, eds. 1998. *Focus on Form in Classroom Second Language Acquisition*. New York, NY: Cambridge University Press.
- 955 Ellis, N.C. 2006. Selective attention and transfer phenomena in L2 acquisition: contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics* 27, no. 2: 164–94.
- Ellis, N.C. 2017. Salience in usage-based SLA. In *Salience in Second Language Acquisition*, ed. S. Gass, P. Spinner and J. Behney, 21–39. New York: Routledge.
- 960 Farley, A.P. and K. McCollam. 2004. Learner readiness and L2 production in Spanish: processability theory on trial. *Estudios de Lingüística Aplicada* 22, no. 40: 47–69.
- García-Amaya, L. 2012. *Second language fluency and cognition: the study of Spanish second language development in an overseas immersion program and an at-home foreign language classroom* (Unpublished doctoral dissertation). Bloomington, Indiana: Indiana University.
- Garza, T.J. 1991. Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals* 24, no. 3: 239–58.
- 965 Gass, S.M., P. Spinner and J. Behney, eds. 2017. *Salience in Second Language Acquisition*. New York: Routledge.
- Geeslin, K. 2003. A comparison of copula choice: native Spanish speakers and advanced learners. *Language Learning* 53, no. 4: 703–64.
- Goldschneider, J.M. and R. DeKeyser. 2001. Explaining the ‘natural order of L2 morpheme acquisition’ in English: a meta-analysis of multiple determinants. *Language Learning* 51, no. 1: 1–50.
- 970 Guasch, M., R. Boada, P. Ferré and R. Sánchez-Casas. 2013. NIM: a web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods* 45, no. 3: 765–71.
- Gudmestad, A. 2012. Acquiring a variable structure: an interlanguage analysis of second-language mood use in Spanish. *Language Learning* 62, no. 2: 373–402.
- Han, Z., E.S. Park and C. Combs. 2008. Textual enhancement of input: issues and possibilities. *Applied Linguistics* 29, no. 4: 597–618.
- Huang, H.C. and D.E. Eskey. 1999. The effects of closed-captioned television on the listening comprehension of intermediate English as a second language (ESL) students. *Journal of Educational Technology Systems* 28, no. 1: 75–96.
- 975 Izura, C., F. Cuetos and M. Brysbaert. 2014. Lextale-Esp: a test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica* 35, no. 1: 49–66.
- Jourdenais, R., M. Ota, S. Stauffer, B. Boyson and C. Doughty. 1995. Does textual enhancement promote noticing? A think-aloud protocol analysis. In *Attention and Awareness in Foreign Language Learning*, ed. R. Schmidt, 182–209. Honolulu: University of Hawaii Press.
- 980 Jung, U.O. 1990. The challenge of broadcast videotex to applied linguistics. *IRAL-International Review of Applied Linguistics in Language Teaching* 28, no. 3: 201–20.
- Kroll, J.F. and E. Stewart. 1994. Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33, no. 2: 149–74.
- Lee, S.K. and H.T. Huang. 2008. Visual input enhancement and grammar learning: a meta-analytic review. *Studies in Second Language Acquisition* 30, no. 3: 307–31.
- 985 Lee, J.F. and P.A. Malovrh. 2009. Linguistic and non-linguistic factors affecting OVS processing of accusative and dative case pronouns by advanced L2 learners of Spanish. In *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, ed. J. Collentine, M. García, B. Lafford and F. Marcos Marín, 105–16. Somerville, MA: Cascadia Proceedings Project.
- Lee, M. and A. Révész. 2018. Promoting grammatical development through textually enhanced captions: an eye-tracking study. *The Modern Language Journal* 102, no. 3: 557–77.
- 990 Leeman, J., I. Arteagoitia, B. Fridman and C. Doughty. 1995. Integrating attention to form with meaning: focus on form in Spanish content-based instruction. In *Attention and Awareness in Foreign Language Learning*, ed. R. Schmidt, Tech. Rep. No. 9, 217–58. Honolulu: University of Hawaii, Second Language Teaching and Curriculum Center.
- Lemhöfer, K. and M. Broersma. 2012. Introducing LextALE: a quick and valid lexical test for advanced learners of English. *Behavior Research Methods* 44, no. 2: 325–43.
- Leow, R.P. and A. Martin. 2017. Enhancing the input to promote salience of the L2: A critical overview. In *Salience in Second Language Acquisition*, ed. S. Gass, P. Spinner and J. Behney, 167–86. New York: Routledge.
- 995 Li, P., F. Zhang, E. Tsai and B. Puls. 2014. Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition* 17, no. 3: 673–80.
- Markham, P.L. 1989. The effects of captioned television videotapes on the listening comprehension of beginning, intermediate, and advanced ESL students. *Educational Technology* 29, no. 10: 38–41.
- Markham, P.L. 1993. Captioned television videotapes: effects of visual support on second language comprehension. *Journal of Educational Technology Systems* 21, no. 3: 183–91.
- 1000 Markham, P.L. 1999. Captioned videotapes and second-language listening word recognition. *Foreign Language Annals* 32, no. 3: 321–28.

- Montero Perez, M., W. Van Den Noortgate and P. Desmet. 2013. Captioned video for L2 listening and vocabulary learning: a meta-analysis. *System* 41, no. 3: 720–39.
- Montero Perez, M., E. Peters, G. Clarebout and P. Desmet. 2014. Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology* 18, no. 1: 118–41.
- 1005 Montero Perez, M., E. Peters and P. Desmet. 2015. Enhancing vocabulary learning through captioned video: an eye-tracking study. *Modern Language Journal* 99, no. 2: 308–28.
- Muñoz, C. 2017. The role of age and proficiency in subtitle reading. An eye-tracking study. *System* 67: 77–86.
- Nation, P. 2001. *Learning Vocabulary in Another Language*. New York: Cambridge University Press.
- Neuman, S.B. and P. Koskinen. 1992. Captioned television as comprehensible input: effects of incidental word learning from context for language minority students. *Reading Research Quarterly* 27, no. 1: 94–106.
- 1010 Overstreet, M. 1998. Text enhancement and content familiarity: the focus of learner attention. *Spanish Applied Linguistics* 2: 229–58.
- Plass, J.L. and L.C. Jones. 2005. Multimedia learning in second language acquisition. In *The Cambridge Handbook of Multimedia Learning*, ed. R.E. Mayer, 467–488. New York: Cambridge University Press.
- Price, K. 1983. Closed-captioned TV: an untapped resource. *MATESOL Newsletter* 12, no. 2: 1–8.
- RStudio Team. 2015. *RStudio: Integrated Development for R*. Boston, MA: RStudio, Inc. URL: <http://www.rstudio.com/>.
- 1015 Schmidt, R. 2001. Attention. In *Cognition and Second Language Instruction*, ed. P. Robinson, 3–32. Cambridge: Cambridge University Press.
- Sharwood Smith, M. 1993. Input enhancement in instructed SLA: theoretical bases. *Studies in Second Language Acquisition* 15, no. 2: 165–79.
- Sydorenko, T. 2010. Modality of input and vocabulary acquisition. *Language Learning & Technology* 14, no. 2: 50–73.
- Tomlin, R.S. and V. Villa. 1994. Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition* 16, no. 2: 183–203.
- 1020 Vanderplank, R. 2010. Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching* 43, no. 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, no. 2: 235–250.
- VanPatten, B. 1987. Classroom learners’ acquisition of *ser* and *estar*: accounting for developmental patterns. In *Foreign Language Learning: A Research Perspective*, ed. B. VanPatten, T. Dvorak and J. Lee, 61–75. Rowley, MA: Newbury House.
- 1025 Vázquez Rozas, V. 2006. Gustar-type verbs. In *Functional Approaches to Spanish Syntax. Lexical Semantics, Discourse and Transitivity*, ed. J. Clancy Clements and J. Yoon, 80–114. New York: Palgrave MacMillan.
- Webb, S. 2008. Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition* 30, no. 1: 79–95.
- Winke, P.M., S. Gass and T. Sydorenko. 2010. The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology* 14, no. 1: 65–86.
- 1030 Winke, P.M., T. Sydorenko and S. Gass. 2013. Factors influencing the use of captions by foreign language learners: an eye-tracking study. *Modern Language Journal* 97, no. 1: 254–75.
- Wulff, S. and N.C. Ellis. 2018. Usage-based approaches to SLA. In *Bilingual Cognition and Language: The State of the Science Across its Subfields*, ed. D. Miller, F. Bayram, J. Rothman and L. Serratrice, 37–56. Amsterdam: John Benjamins.

1035

1040

1045

1050