

The Effects of Textually Enhanced Captions on Written Elicited Imitation in L2 Grammar

LORENZO GARCÍA-AMAYA¹ and MYRNA C. CINTRÓN-VALENTÍN²

¹ *The University of Michigan, Romance Languages and Literatures, 4108 Modern Languages Building, 812 East Washington Street, Ann Arbor, Michigan 48109-1275 Email: lgarciaa@umich.edu*

² *State of California Department of Justice, Research Center, 4949 Broadway, Room G-105, Sacramento, CA, 95820 Email: myrna.cintronvalentin@gmail.com*

<ABSTRACT>

As second language (L2) research adapts to the dynamics of multimodal instruction, researchers are exploring the role of captioning and textual enhancement (TE) on learner development. The present experiment assesses the differential effects of textually enhanced captions on learners' elicited imitation of L2 Spanish grammar. Twenty-eight English-speaking intermediate L2 Spanish learners saw 3 videos that focused on *gustar*-type verbs, the preterite–imperfect contrast, or the subjunctive, respectively. Each video included 3 conditions: sentences without captions (NC); sentences with target verbs highlighted entirely (TE1); or sentences with highlighting on inflectional morphemes (TE2). Using a novel written elicited imitation task, we uncovered consistent positive effects of both TE conditions compared to NC, with an added advantage of TE2 for the subjunctive. The findings demonstrate that (a) TE with captions contribute to increased L2 accuracy, (b) the elicited imitation of some grammatical structures is more conducive to enhancement than others, and (c) there is space for future investigation into the factors that mediate the effectiveness of multimodal interventions.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/modl.12740](https://doi.org/10.1111/modl.12740).

This article is protected by copyright. All rights reserved.

<END ABSTRACT>

Keywords: textual enhancement; captioning; written elicited imitation; salience; noticing; grammar acquisition; Spanish

The acquisition of grammar is one of the most challenging aspects of second-language (L2) learning (DeKeyser, 2005). The second language acquisition (SLA) literature provides various accounts explaining this difficulty, including age effects (e.g., Johnson & Newport, 1989), individual learner-based differences (e.g., Dörnyei, 2005), and input-processing differences between native (L1) speakers and L2 learners (e.g., VanPatten, 1996). The linguistic features of individual target grammatical items can likewise contribute to acquisitional challenges: frequency (e.g., Ellis, 2002), complexity (Gebhart, Newport, & Aslin, 2009), and, critically, perceptual salience (e.g., Goldschneider & DeKeyser, 2001). For example, low perceptual salience of certain grammatical features, such as inflectional suffixes, contributes to L2 learners' difficulties in acquiring them (Cintrón-Valentín & Ellis, 2016; Gass, Spinner, & Behney, 2018; Goldschneider & DeKeyser, 2001). The challenges posed by perceptual salience relate to one key question that shapes much of the L2 literature—namely, whether learner attention can be enhanced toward commonly unattended input features (e.g., Gass et al., 2018; Goldschneider & DeKeyser, 2001).

To mitigate the acquisitional challenges presented by low perceptual salience, SLA researchers have explored the role of form-focused-instruction (FFI) techniques, such as textual enhancement (TE) and explicit grammar instruction, in rendering target structures

more salient (Norris & Ortega, 2000). Critically, TE is typically limited to unimodal mediums that focus on the enhancement of grammatical cues through written input in the absence of pictorial or aural cues. However, with the increased availability of multimedia language-learning materials, FFI research can more deeply scrutinize the role of multimodal input (i.e., aural, written, and visual) in facilitating L2 development (e.g., Blake, 2013; Plass & Jones, 2005).

One promising multimodal technique is captioned video (e.g., Ghia, 2012; Montero Perez, Van Den Noortgate, & Desmet, 2013; Vanderplank, 2010). Salience-raising in captioned video presents a unique experimental opportunity in SLA, creating space to augment the existing literature that documents the positive effects of captioned media on L2 comprehension and vocabulary learning (e.g., Muñoz, 2017; Winke, Gass, & Sydorenko, 2013). Even within the realm of multimedia captioning in grammar learning, the extent of any positive effects are understood at a very general level (e.g., Cintrón-Valentín, García-Amaya, & Ellis, 2019; Lee & Révész, 2018). For example, captioned media may not be reliably effective for all grammar structures, indicating that greater or more nuanced instructional support may be needed for certain structures for learners to fully benefit from multimodal input.

Weighing these considerations, we designed an innovative methodology to examine the differential effects of TE alongside captions in a multimedia setting. Specifically, we compare the effects of tailoring TE on a full lexical entry to those of tailoring TE on a target morpheme only. Rather than focus on textual media alone, our study incorporates animated videos to provide an auditory learning channel while also implementing dynamic visual input. Using this multimodal design, we analyze learners' elicited imitation (EI; Erlam, 2006)

of three grammatical structures in L2 Spanish: *gustar*-type verbs, the preterite–imperfect contrast, and the subjunctive in noun clauses. Our experimental design addresses the need to adequately measure learners’ immediate detection of perceptually enhanced input (Han, Park, & Combs, 2008), while further augmenting the developmental literature that includes pre–posttest designs (e.g., Cintrón–Valentín et al., 2019).

<A>BACKGROUND

Saliency, Grammar Learning, and Elicited Imitation in an L2

The role of saliency as it relates to the perceptual distinctiveness of a linguistic cue in their input has received increasing interest in recent years, such as from Ellis (2017): “Salient items or features are attended, are more likely to be perceived, and are more likely to enter into subsequent cognitive processing and learning” (p. 21; see also Gass et al., 2018). This is especially relevant for the acquisition of grammar given the low perceptual saliency that characterizes certain inflectional morphemes (e.g., Goldschneider & DeKeyser, 2001). In fact, one common observation in SLA is that despite the vast availability of grammatical forms in their input, L2 learners often ignore specific aspects of morphological structure and focus their attentional resources to the meanings of open-class words during input processing (e.g., VanPatten, 1996).

One way of counteracting the low saliency of grammatical forms is to provide learners with enhanced input designed to render target structures more perceptually distinct (Sharwood Smith, 1993). TE involves visual manipulations in written input (e.g., bolding, underlining, capitalization) that facilitate learners’ processing of target grammatical forms. For example, LaBrozzi (2016) showed that increased font size on L2 Spanish aspectual

morphemes led to greater recognition of present versus preterite morphemes than in a control condition and in a capital-letter manipulation (e.g., Jourdenais et al., 1995; Leeman et al., 1995; Overstreet, 1998). Despite such advancements, the experimental research in this area yields largely inconsistent findings (e.g., Han et al., 2008; Lee & Huang, 2008), with some research suggesting that the efficacy of TE may be modulated by the linguistic form in question (e.g., Comeaux & McDonald, 2017). Chiefly, most TE studies compare the effects of different enhancement manipulations (e.g., normal vs. underlined text) but do not consider tailoring the TE to a target morpheme in comparison to TE on a full lexical entry. As pointed out in Lee & Huang (2008), the next step in exploring any substantive effect of TE is through the design of studies that explore how learners respond to enhanced forms and whether the processing of such enhanced forms promotes L2 grammar acquisition. In our view, this step involves a more focused analysis of the effects of TE on a target morpheme, emphasizing the appropriate inflectional and functional considerations.

In the current study, we aim to assess which TE designs best focus learner attention on L2 grammar by assessing participants' immediate reproduction of enhanced target forms. We conceptualize learners' immediate reproduction of a grammar form as their EI (Erlam, 2006) in response to experimental variations of enhanced captioning. The EI method, in which participants are instructed to reproduce input from the L2, is commonly used in SLA as a measure of language proficiency (Bowden, 2016; Tracy–Ventura et al., 2014). One key assumption underlying this method is that learners should be more successful in reproducing utterances that contain familiar grammatical structures and less successful when the grammatical structures exceed their capacities (Yan et al., 2016). In our study, we

developed a written adaption of the EI design to determine whether TE draws learners' attention to relevant parts of the input, thereby facilitating their reproduction.

Notably, only a handful of studies have investigated immediate learners' noticing while being exposed to enhanced input (e.g., Cintrón-Valentín & Ellis, 2015; Indrarathne & Kormos, 2017; Winke, 2013). These studies have included eye-tracking methods that measure learners' visual attention to form, capturing noticing as it unfolds (see Roberts & Siyanova-Chanturia, 2013). Through a novel EI paradigm designed for this study, in which learners reproduce input forms in written format, we add to this literature by measuring learners' accuracy of L2 Spanish grammar immediately after receiving enhanced written input in a multimodal setting. In the next section, we present an overview of the captioning research as it relates to multimodal learning and its potential for promoting L2 noticing.

Multimedia SLA, Captioning, and Noticing

As pointed out in Han et al. (2008) and Lee & Huang (2008), TE research has yielded inconsistent results regarding L2 learning—one possible explanation might be that most experimental designs rely on unimodal media inputs only (i.e., written input only). Theorizing this point, Mayer's (2003) cognitive theory of multimedia learning proposes that multimedia input yields stronger benefits for L2 learning than input derived from a single medium. Per Plass & Jones (2005), SLA multimedia is defined as input in the form of words and images that promotes meaningful output and target-language interaction. In particular, captioned media is a commonly used tool among L2 instructors, with advances in multimedia technology making these resources more accessible in L2 classrooms (e.g., Blake, 2013; Ghia, 2012; Plass & Jones, 2005).

Winke, Gass, & Sydorenko (2010) attributed the usefulness of captioned media to matters of attention, suggesting that this medium draws learners' attentional focus to unknown forms, thereby promoting noticing and subsequent learning through repeated exposure. This view aligns with foundational models of SLA centered on the role that attention plays in facilitating successful acquisition (e.g., Gass et al., 2018; Schmidt, 2001; 2010). Schmidt's (2001) noticing hypothesis (2001), for instance, holds that conscious attention to linguistic forms in the input is an important precondition to learning (but see Tomlin & Villa, 1994). The conceptualization of noticing, compared to understanding (or awareness), is further delineated in Schmidt (2010). Specifically, Schmidt distinguished 'noticing,' or the conscious registration of attended linguistic input, from 'understanding,' which is the knowledge of metalinguistic rules. On the one hand, a learner can consciously notice a particular language structure; on the other, they may not have the ability to understand or apply the particular structure, or to generalize its underlying rules, in novel contexts (for a detailed discussion of these terms, see Schmidt, 2001, 2010). Clearly, one key aim of TE research is to promote the noticing of target linguistic items through enhanced input, thereby facilitating subsequent learning through awareness and understanding (e.g., Han et al., 2008; Sharwood Smith, 1993). As a motivating point for this study, we thus propose that the use of captioning can help learners notice L2 grammatical features (per Schmidt, 2001, 2010) in a modern learning environment that employs multimedia tools (Mayer, 2003).

Regarding studies published on captioning effects in an L2 setting, the early research primarily focused on determining if captioned video was more efficacious than noncaptioned video in (a) improving learner comprehension of video content (e.g.,

Markham, 1999), and (b) promoting vocabulary learning (e.g., Huang & Eskey, 1999).

Although vocabulary learning and comprehension have remained central in the L2 captioning literature (e.g., Montero Perez et al., 2013; Tragant Mestres & Pellicer-Sánchez, 2019; Vanderplank, 2010), recent work has explored the effect of specific experimental factors that mediate the effectiveness of captioning, such as ordering effects (Winke et al., 2010), modality effects (Sydorenko, 2010), and learner variables, including age and proficiency (Muñoz, 2017). Thus far, the research focused on captioning in grammar learning is more limited in comparison. One preliminary outcome in this area is that captioned media may not be reliably effective for all grammar structures (Lee & Revesz, 2018) and certain structures may require greater instructional support to fully benefit from multimodal input (Cintrón-Valentín et al., 2019). We aim to augment this literature by examining learners' written reproductions of three Spanish grammar constructions of varying syntactic and discourse dependence, as described in the next section.

<A>THE CURRENT STUDY

Target Structures

One fundamental question underlying research on L2 learning and processing is whether learners are able to track relationships between words and phrases in discourse (e.g., Vuong, Mayer, & Christiansen, 2015; Wilson et al., 2018). Embedded in L2 grammar acquisition is the learning of morphosyntactic constituents and their dependencies. For instance, grammar structures that depend on a relationship between nonadjacent words are psycholinguistically taxing for L2 learners due to the distance required to process the relationship (Vuong et al., 2015; Wilson et al., 2018).

Simultaneously, research shows that pedagogical interventions designed to enhance the salience of nonadjacent forms can facilitate their learning (e.g., Newport & Aslin, 2004). In this study, we analyze the effect of captions with TE on the reproduction of three structures of varying syntactic and discourse dependence: *gustar*-type verbs, in which the related elements are not separated by intervening material; the preterite–imperfect contrast, which involves more complex grammatical relations than *gustar*-type verbs due to the need to track the surrounding discourse context (cf. Bardovi–Harlig, 1998); and the subjunctive mood, with nonadjacent morphosyntactic dependencies. The following paragraphs offer a brief overview of each structure.

<C>*Gustar-Type Verbs*. The Spanish verb *gustar* is often translated as ‘to like’ in English. Whereas English *like* codes the experiencer of the action as the subject and the liked stimulus as the direct object, in Spanish, *gustar* codes the experiencer as an indirect object and the stimulus as the sentential subject (i.e., ‘it is pleasing to me;’ see Vázquez Rosas, 2006). Previous literature on the acquisition of *gustar*-type verbs examines the processing of the clitic pronoun preceding the verb (e.g., Lee & Malovrh, 2009). We focus here on an additional challenge, namely the agreement between verb morphology and its subject. Specifically, the conjugated verb in Spanish depends on whether the liked entity is singular or plural, as in *me gusta la manzana* ‘I like the apple’ compared to *me gustan las manzanas* ‘I like apples.’

<C>*The Preterite–Imperfect Contrast*. The standard usage of the Spanish past-tense system requires that learners understand the aspectual distinction between the preterite and imperfect (Liskin–Gasparro, 2000). Preterite forms characterize past actions as having a definitive beginning and endpoint (e.g., *caminé* ‘I walked’), whereas imperfect forms

characterize past actions or states as in progress (e.g., *caminaba* ‘I was walking / I used to walk’). Understanding the preterite–imperfect contrast within specific semantic or discourse contexts presents a considerable challenge for SLA (Bardovi–Harlig, 1998; Overstreet, 1998). Additionally, tense–aspect morphological forms such as the preterite and imperfect differ in their frequency distribution in Spanish, thereby reducing L2 learners’ exposure to their direct contrast (Liskin–Gasparro, 2000). As a result, Blyth (2005) asserted that pedagogical interventions that render surface forms more frequent and salient can allow learners to focus on form more meaningfully. This latter assertion serves as a motivating point for our study.

<C>*Subjunctive in Noun Clauses.* The Spanish subjunctive mood is typically used in sentences involving subordination, wherein the subject of the main clause exerts influence or will over the subject of the subordinate clause—in this case, a noun clause that serves as the object of the verb (Gudmestad, 2012). The subjunctive is often described as a late-emerging structure for both L1 and L2 acquisition given its low frequency and the low perceptual salience of the subjunctive inflection (DeKeyser & Prieto Botana, 2013). However, research has shown that breaking down the morphosyntactic components of this construction can facilitate its acquisition, regardless of learners’ readiness (Collentine, 2013). To this end, in the current study, both the main-clause verb, which acts as a cue to the subjunctive, and the subordinated subjunctive verb were textually enhanced. We will explore the effects of such enhancements through the research questions listed in the following section.

Research Questions and Hypotheses

This study aimed to extend previous research on TE, captioning, and L2 grammar by exploring three research questions through a written elicited imitation (WEI) experimental paradigm.

- RQ1. What is the relative effect of lexical-item-enhanced TE (TE1), morpheme-enhanced TE (TE2), or no TE on the WEI of *gustar*-type verbs in L2 Spanish?
- RQ2. What is the relative effect of lexical-item-enhanced TE (TE1), morpheme-enhanced TE (TE2), or no TE on the WEI of the preterite–imperfect contrast in L2 Spanish?
- RQ3. What is the relative effect of lexical-item and dependency-enhanced TE (TE1), morpheme- and dependency-enhanced TE (TE2), or no TE on the WEI of the subjunctive in noun clauses in L2 Spanish?

We investigated the effects of TE within the captioning line on L2 WEI through three experimental conditions: the no-captions (NC) condition, which presented L2 audio but no material in the captioning line; the TE1 condition, which highlighted verbs in their entirety; and the TE2 condition, which highlighted only the critical morphological and grammatical cues. The NC condition, without captions or enhancement, served as a baseline measure of participants' prior knowledge of each structure. The integration of a NC condition further allows us to investigate the extent to which any group-level effects of the experimental conditions (i.e., TE1 and TE2) remain consistent across individual learners (cf. Larsen–Freeman, 2018). Our within-subjects design (see 'Methods' section) allows for such comparisons at the level of the individual learner.

Two hypotheses guided our research.

- H1. For all target structures, both TE conditions will render more accurate reproductions than the NC condition.
- H2. The preterite–imperfect contrast and the subjunctive will receive an added benefit from TE2 beyond that of TE1.

Given the usefulness of TE and captioned media for directing learners' attention to L2 input (e.g., Winke et al., 2010), WEI should benefit from TE. The preterite–imperfect contrast and the subjunctive are context-dependent structures defined by their surrounding discourse or morphosyntactic context; previous research shows that enhancing the salience of such long-distance grammar relations can facilitate their learning (e.g., Newport & Aslin, 2004). We thus expect a greater benefit of TE2 over TE1 for the preterite–imperfect and subjunctive structures, but not for *gustar*-type verbs, whose target dependencies were adjacent in our design.

<A>METHODS

Participants

A total of 31 English-speaking L2 learners of Spanish were recruited from two summer-term Spanish courses at a large Midwestern university in the United States. The learners participated in the experiment for credit as one of their course requirements. The average age of all learners was 19.66 ($SD = 0.79$, range = 18 to 21); there were 26 female and 5 male participants. Of the 31 participants, 3 female learners were subsequently excluded from the study because they either were raised bilingually or had recently completed a study-abroad program lasting 2 or more months (see Online Supporting Information A). All participants were sixth-semester intermediate learners of Spanish, having previously taken

their university's grammar-review course required for the Spanish concentration. This grammar course bridges their university's elementary-language program with their upper division courses geared toward Spanish concentrators; all participants therefore had prior instructional experience with the target structures.

Language History Questionnaire and Spanish Grammar Proficiency Test

L2 learners' previous knowledge of target linguistic forms is an important factor when considering the effectiveness of TE on L2 development, in addition to task demands (e.g., Winke, 2013). As mentioned earlier, the NC condition in the experiment (i.e., without captions or TE) served as a baseline measure of participants' prior knowledge of each structure from which to gauge effects of TE1 and TE2. To further gauge prior knowledge, all participants completed a 45-item grammar proficiency test (García-Amaya, 2012) which consisted of a short passage with a series of multiple-choice options covering a broad variety of grammatical items.¹ Participants received 1 point for each correct response, for a maximum of 45 points. We included the results of this test as a control variable in our statistical modeling, thereby controlling for prior knowledge.

The learners also completed a language history questionnaire (LHQ; Li et al., 2014), which included demographic questions in addition to detailed questions about previous language experience. The learners completed the LHQ during the first week of the term and the grammar proficiency test on their third day of the term.

Captioning Experiment

<C>Target Structures. We targeted three grammatical constructions: *gustar*-type verbs, the preterite–imperfect contrast, and the subjunctive in noun clauses. For *gustar*-

type constructions, we used both singular and plural subjects. For preterite and imperfect verbs, we targeted three environments in which each structure can occur—for the preterite, these were single occurrences, precise actions, and consecutive events; for the imperfect, they were habitual occurrences, imprecise actions, and simultaneous events (see Tables C1–C2 in the Online Supporting Information). For the subjunctive, we targeted five constructions: impersonal observations; recommendations; expressions of emotion; doubt; and wishes, desires, and imperatives (see Tables C3–C4 in the Online Supporting Information).

<C>*Animated Videos.* We devised three original videos, one per target grammar structure; within each video, the target structures were presented using all three condition types (NC, TE1, and TE2). Tailoring each video to a specific target structure allowed us to control the placement, randomization, and frequency of occurrence of each grammar item (i.e., 28 preterite or imperfect verbs, and 24 each of subjunctive and *gustar*-type verbs). The video design included generating scripts, recording the characters' voices, and finally animating these scripts (see Online Supporting Information D for excerpts from each of the three scripts).

We created each of the three animated videos using Nawmal (www.nawmal.com), an animation program that allows users to create videos from a menu of predesigned characters and sets. This software allows for the uploading of user-recorded voices directly into the application. In our case, these were the recordings made by the two authors of this study (the male voice by a native speaker of Peninsular Spanish and the female voice by a native speaker of Puerto Rican Spanish), which were then automatically lip-synched to fictional characters. The Nawmal software also supports the inclusion of gestures as the

characters proceed throughout the dialogue, plus camera movements, which help make the scenes feel more dynamic.

All target sentences were visually presented, one at a time, between square brackets to signal that these sentences would need to be recalled for the subsequent written imitation. There were three possible conditions for each target sentence: NC sentences that did not show any text for the target sentences, except the square brackets; TE1 sentences that included target verbs highlighted in their entirety within the target sentences (for the preterite–imperfect contrast and the subjunctive, this implied highlighting the two relevant verbs in addition to their connecting conjunction when applicable); and TE2 sentences, in which only the critical morphological and grammatical cues, and their relations, were highlighted. All captions were added using SRT Edit Pro (www.finalsub.com/sep.html), which facilitated the inclusion of color-coding and bold-facing.

Table 1 offers a summary of the TE1 and TE2 manipulations, and Figures 1 and 2 shows screen captures of two sample manipulations. No distractor items were included for this study.

<INSERT TABLE 1 ABOUT HERE>

TABLE 1

Summary of Textual Enhancement (TE) Manipulations per Grammar Structure

Grammar structure	TE1: Lexical Item (Verb)	TE2: Inflectional Morphemes
<i>Gustar</i> -type verbs	<i>Me molestan los ruidos.</i>	<i>Me <u>molestan</u> los ruidos.</i>
	‘Noises bother me.’	

Preterite–imperfect	<i>Cuando salí de la casa llovía.</i>	<i>Cuando <u>salí</u> de la casa <u>llovía</u>.</i>
	‘When I left home it was raining.’	
Subjunctive in noun clauses	<i>Dudo que pueda venir.</i>	<i>Dudo → que <u> pueda </u>venir.</i>
	‘I doubt that he/she will come.’	

We adopted a within-subjects design, in which all participants saw all three conditions for each of the target grammar topics. The within-subjects design has two advantages: First, it leads to a higher number of participants per condition and second, it allows for within-subject comparisons between the experimental conditions. We further created three orders for each video so that the target sentences would not be repeated in the same condition for all learners. For each of the three orders, the presentation of each of the three conditions was randomized (see Online Supporting Information B).

<INSERT FIGURE 1 ABOUT HERE>

FIGURE 1

Screenshot of Lexical-Item Textual Enhancement (TE1) for the Subjunctive in Noun Clauses



Note. English translation: 'I recommend that you be careful, and think before you speak.'
The main clause verb, conjunction *que* 'that,' and subordinate subjunctive verb were in bold and yellow.

<INSERT FIGURE 2 ABOUT HERE>

FIGURE 2

Screenshot of Morpheme Textual Enhancement (TE2) for the Subjunctive in Noun Clauses



Note. English translation: 'I recommend that you be careful, and think before you speak.' The main clause verb was in bold and orange, followed by an orange arrow. The subjunctive verb was underlined; the target subjunctive morpheme was in bold and yellow.

<C>Written Elicited Imitation. We developed an adapted version of the EI task to assess the effect of TE on learners' abilities to reproduce grammatical structures when prompted. As an innovation, we developed a WEI task to investigate whether TE draws learners' attention to relevant parts of the input through the reproduction of target grammatical structures in a written format (see Vinther, 2002). Through our design, learners were informed that the majority of the videos would include Spanish-language captioning, which consisted of white, nonbolded text on a black background on the bottom of the screen, superimposed over the video image. At unpredictable points in the videos, square brackets appeared in the captioning line; once the audio of the spoken sentence ended, the

video paused for up to 20 seconds. Learners knew that during these 20 seconds, they had to type, verbatim, all of the words they could remember from the target sentence.

As noted in Tracy–Ventura et al. (2014), there is a concern that short, immediately-repeated stimuli may be automatically imitated instead of processed. To avoid participants' automatic imitation of the stimuli, we incorporated sentences ranging from 9–44 syllables with a 3-second pause (see Tracy–Ventura et al., 2014, for further elaboration). To further minimize such imitation, we included an average 4.67 syllables prior to each target verb (range 0–16).

The single experimental session took place on the second day of the summer program in a large auditorium where each participant was provided with a laptop and headset. The full experiment was programmed in Open Sesame (Mathôt, Schreij, & Theeuwes, 2012) and took less than 1 hour to complete. Participants were presented with the three animated videos (one per target structure) in random order. To alleviate concerns that the appearance of nontarget structures would distract from the target grammar points (e.g., Robinson, 2003), we focused our experimental design on learners who had previous classroom exposure to all three structures. The participants' prior exposure to the target structures suggests that the task demands were not as great as if they had received no prior instruction focused on these items.

Data Analysis

For each trial, we measured the accuracy of the written target-grammar verbs from each sentence. Each response received a score of 1 or 0 based on the usage of the morphological ending (1 = correct usage; 0 = incorrect usage), which was computed via an

Excel formula programmed to identify the appropriate target morpheme from each sentence. This binary outcome served as the dependent measure in the regression model.

The statistical analysis was conducted using RStudio version 1.0.143 (RStudio Team, 2015). The data were analyzed through a multilevel logistic linear regression model utilizing the “glmer()” function within the “lme4” package for R (Bates et al., 2015). The independent variables included the interaction of grammar structure (*gustar*-type verbs, preterite–imperfect, and subjunctive) and captioning (NC, TE1, and TE2), in addition to each of the individual predictor terms. We added grammar proficiency, syllables of target sentence (range 9–44), syllables to target verb (range 0–16), and syllables after target verb (range 0–21) as control variables. The model also included random intercept terms for subject and sentence. For model diagnostics, we checked for collinearity in the predictor variables and also investigated residual distribution plots, Q-Q plots, and plots of residual values versus fitted values.

Since our design focused on differences between each captioning condition within each grammar construction, we tested the overall significance of the Captioning × Grammar Structure interaction. To focus on the previously stated hypotheses, we refit the same model multiple times using different reference levels for the two predictors. We report the corresponding β coefficients and their standard errors, p values, and, finally, odds ratios (OR) as a measure of effect size (Plonsky et al., 2014). An OR greater than 1 means that for the tested condition, there are greater odds of obtaining a higher score than under the reference condition, whereas an OR of less than 1 means that the reference condition has higher odds for obtaining a correct score than the tested condition. Finally, we set the significance level to $\alpha = 0.05$ for all tests.

<A>RESULTS

The ANOVA output generated by the linear regression model is listed in Table 2. The type III *F*-tests returned a significant effect of the interaction of interest (Captioning × Grammar Structure) in addition to two control variables (grammar proficiency and syllables after target verb). Due to the significant effect of the targeted interaction, we will not draw inferences based on the main effects of captioning and grammar structure.

<INSERT TABLE 2 ABOUT HERE>

TABLE 2

ANOVA Output of Model of Interest

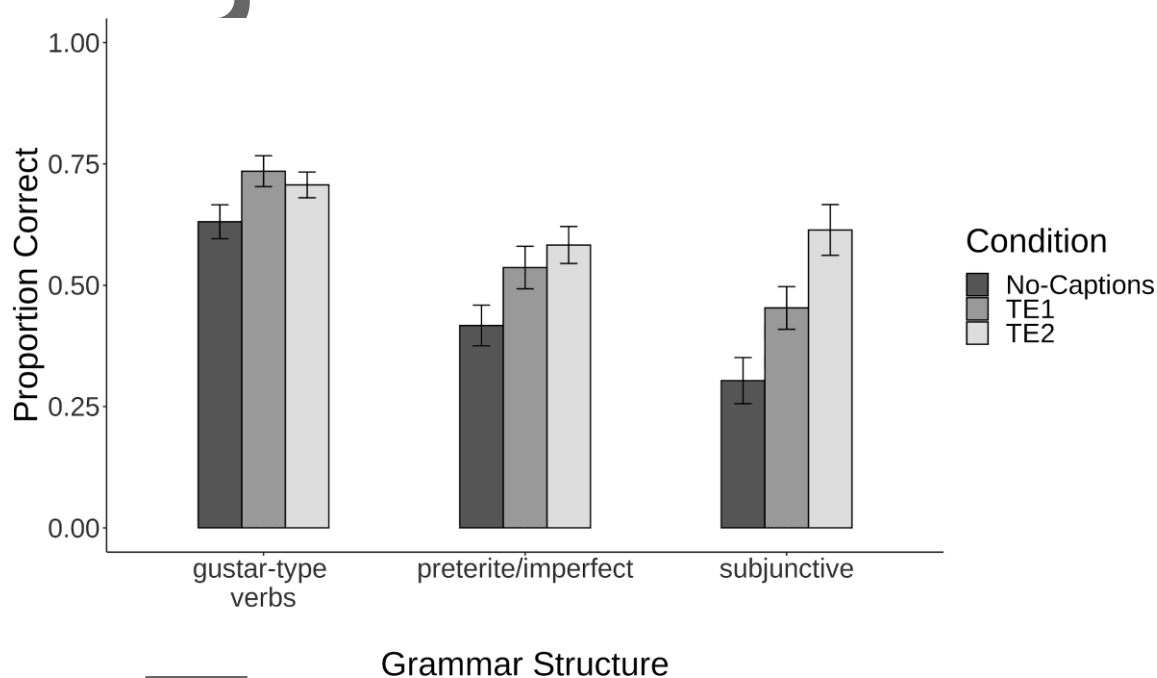
Variable	Degrees of Freedom	<i>F</i> value	<i>p</i> value
Captioning × Grammar Structure	4	4.018	0.003
Captioning	2	30.755	< 0.001
Grammar structure	2	5.562	0.004
Grammar proficiency	1	20.415	< 0.001
Syllables of target sentence	1	3.635	0.057
Syllables to target verb	1	0.074	0.786
Syllables after target verb	1	10.005	0.002

Figure 3 plots the accuracy proportion scores for each grammar topic per captioning condition, and Table 3 incorporates the corresponding descriptive data. In the following subsections, we outline the major statistical findings per grammar construction.

<INSERT FIGURE 3 ABOUT HERE>

FIGURE 3

Proportion Correct Scores for All Structures by Condition



Note. TE1 = lexical-item textual enhancement; TE2 = morpheme textual enhancement. Error bars represent two standard errors.

<INSERT TABLE 3 ABOUT HERE>

TABLE 3

Accuracy Proportion Scores for Each Grammar Topic by Condition

Condition	Mean	SD	95% Confidence Interval
-----------	------	----	-------------------------

Accuracy scores for *gustar*-type verbs

NC	0.63	0.19	[0.56, 0.69]
TE1	0.74	0.16	[0.68, 0.80]
TE2	0.70	0.15	[0.65, 0.75]

Accuracy scores for the preterite–imperfect contrast

NC	0.42	0.22	[0.33, 0.50]
TE1	0.54	0.23	[0.45, 0.62]
TE2	0.58	0.20	[0.51, 0.68]

Accuracy scores for the subjunctive in noun clauses

NC	0.30	0.25	[0.21, 0.40]
TE1	0.45	0.23	[0.37, 0.54]
TE2	0.61	0.28	[0.51, 0.72]

Note. NC = no captions; TE1 = lexical-item textual enhancement; TE2 = morpheme textual enhancement.

Gustar-Type Verbs

The data for *gustar*-type verbs are plotted in the left panel of Figure 3 and summarized in the first three rows of Table 3. The pattern for this construction suggests an advantage of both captioning conditions (i.e., TE1 and TE2) over NC.

To investigate the effects of captioning on accuracy, we fit a generalized linear mixed-effects model, which included the factors mentioned in the previous section. The first comparison, with *gustar*-type verbs and the NC condition as the reference levels, revealed a significant positive effect for TE1, $\beta = 0.878$, $SE = 0.270$, $p = 0.001$, $OR = 2.406$. The OR indicates that for the TE1 stimuli, the odds of obtaining a correct score are approximately 140% higher than for the NC stimuli. The effect for TE2 was also in the expected direction, with the model yielding values that approached significance, $\beta = 0.483$, $SE = 0.265$, $p = 0.068$, $OR = 1.621$. Regarding differences between the two captioning conditions, the model did not reveal a significant effect when comparing TE2 to TE1, $\beta = -0.394$, $SE = 0.276$, $p = 0.153$, $OR = 0.674$.

The Preterite–Imperfect Contrast

The middle panel of Figure 3 plots the preterite–imperfect accuracy data (see also the middle three rows of Table 3) and reveals a general advantage of both TE conditions over NC. We find positive significant differences between the two treatments and NC: TE1, $\beta = 0.851$, $SE = 0.222$, $p = 0.001$, $OR = 2.342$; TE2, $\beta = 1.086$, $SE = 0.222$, $p < 0.001$, $OR = 2.962$. That is, both captioned conditions led to greater reproduction accuracy than NC. However, there was no significant difference between TE1 and TE2, $\beta = 0.235$, $SE = 0.220$, $p = 0.286$, $OR = 1.265$.

The Subjunctive in Noun Clauses

The results for the subjunctive in noun clauses are plotted in the right panel of Figure 3 (see also last three rows of Table 3). There appears to be an advantage of both captioning conditions over the NC condition, with an added advantage of TE2 over TE1.

The results for the t tests in the regression model returned significant positive effects of TE1, $\beta = 0.883$, $SE = 0.256$, $p < 0.001$, $OR = 2.418$; and TE2, $\beta = 1.778$, $SE = 0.263$, $p < 0.001$, $OR = 5.918$, over NC. Additionally, the comparison between TE1 and TE2 revealed a significant positive effect for TE2, $\beta = 0.895$, $SE = 0.253$, $p < 0.001$, $OR = 2.447$, confirming our observation that there is an added advantage for the TE2 manipulation.

Individual Data

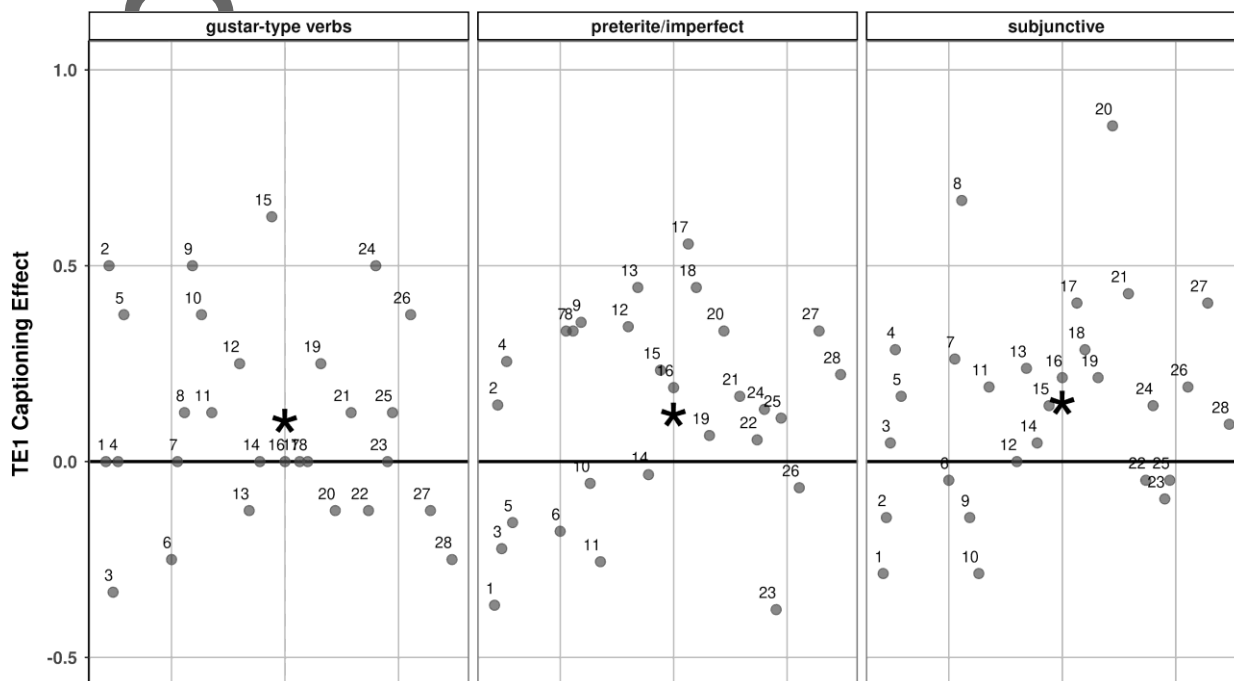
To determine the extent to which group-level effects are present across all learners (cf.

Muñoz, 2017), we calculated individual per-structure captioning-effect scores for each participant (see Figures 4 and 5). 'Captioning Effect' on the y axes of Figures 4 and 5 represents the calculated difference between each participant's TE1 and NC mean values and between each participant's TE2 and NC mean values, respectively. In both figures, a score above 0 indicates a positive effect for the respective TE condition, when compared to NC, and a score below 0 indicates a negative effect. All participant means are labeled, facilitating a within-subjects comparison based on the experimental conditions (see Online Supporting Information E).

<INSERT FIGURE 4 ABOUT HERE>

FIGURE 4

Lexical-Item Textual Enhancement (TE1) Captioning-Effect Scores

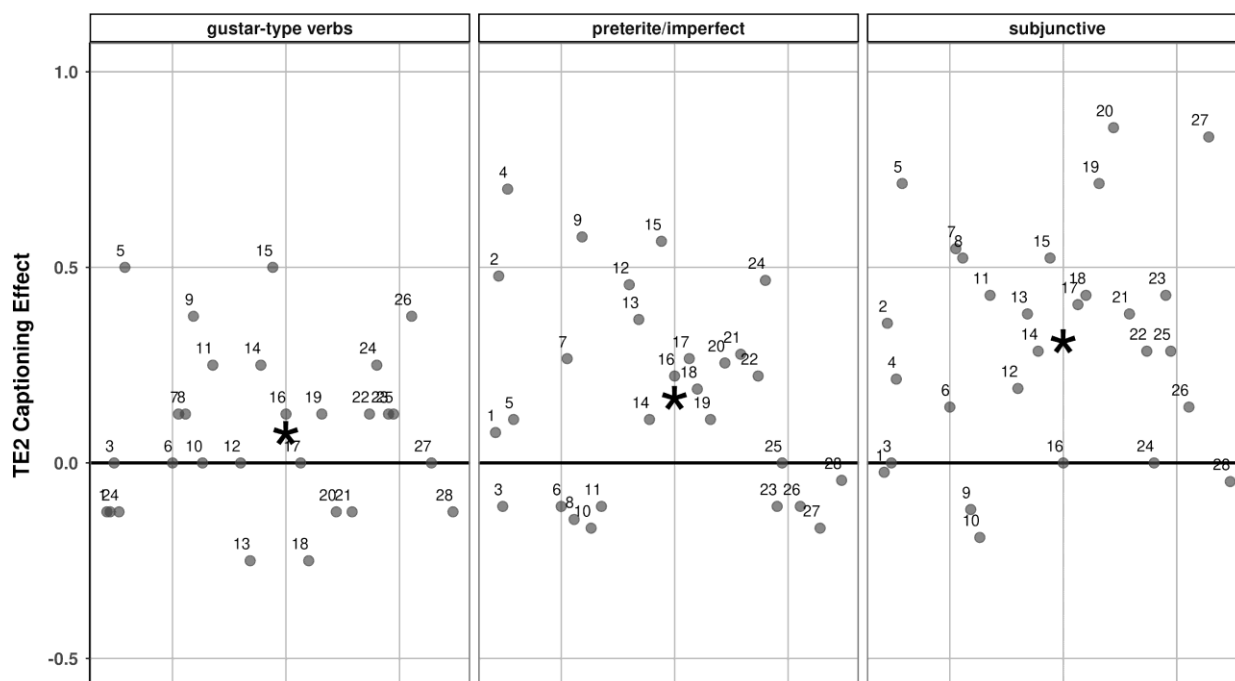


Note. Scores represent the calculated difference between each participant's TE1 and no-captioning (NC) mean values. The data points are spread apart along x axes for readability purposes only. Each black asterisk represents the per-structure mean.

<INSERT FIGURE 5 ABOUT HERE>

FIGURE 5

Inflectional Morpheme Textual Enhancement (TE2) Captioning-Effect Scores



Note. Scores represent the calculated difference between each participant's TE2 and no-captioning (NC) mean values. The data points are spread apart along x axes for readability purposes only. Each black asterisk represents the per-structure mean.

The overall pattern visualized by the means of the TE1 and TE2 effects are similar to the group-level findings reported previously, with broadly positive TE effects for all structures and additional TE2 effects (compared to TE1) for the subjunctive. Further, all learners demonstrated sensitivity to TE (i.e., scores above 0) for at least one structure, and especially to TE2 in the subjunctive items (21 of the 28 learners). Two participants (15 and 19) consistently obtained positive values for all grammar structures in both TE conditions. At the same time, we observe some heterogeneity in the captioning-effect scores, with some structures and TE conditions showing scores at or below 0—for example, for *gustar-type* verbs, 19 of the 28 learners showed no effect or a negative effect for at least one of the two TE manipulations. Additionally, four participants (1, 6, 10, and 28) returned negative scores for the same structure in both TE

conditions. Thus, while TE effects are positive at the group level, individual scores reveal that not all learners benefitted from the TE manipulations across all structures.

<A>DISCUSSION

In this study, we considered the differential effects of TE on reproducing L2 grammar features using two enhancement methods: TE1, highlighting the entire target word, and TE2, highlighting key morphemes related to the target structure as well as grammatical dependencies. Overall, captions incorporating some type of TE led to increased accuracy in learners' immediate reproductions of the target grammatical forms relative to the NC condition. Our results thus suggest that the provision of the TE led to greater noticing, or initial detection, of the target grammatical features.

RQ1: *Effects of Textual Enhancement on the Written Elicited Imitation of Gustar-Type Verbs*

Regarding *gustar*-type verbs, the statistical results revealed a significant effect of the TE1 condition over NC and an effect approaching significance (in the expected direction) of TE2 over NC. These results provide support for the positive effects of TE on *gustar*-type verbs, corroborating developmental research focused on the L2 learning of this same construction (e.g., Cintrón–Valentín & García–Amaya, 2021; Cintrón–Valentín et al., 2019). Whereas most previous literature on the acquisition of *gustar*-type verbs focuses on the processing of the clitic pronoun preceding the verb (e.g., Lee & Malovrh, 2009), in our study we explored an additional, sometimes overlooked challenge in the acquisition of this construction—namely, the morphological agreement between verb and subject. We

showed that by including TE in multimodal videos, learners can overcome this challenge during L2 reproduction.

We did not uncover statistical differences between the two captioned conditions. One possible explanation for why there was no added effect of morpheme-enhanced TE over lexical-item-enhanced TE might relate to the nature of the syntactic dependencies in question (cf. Vuong et al., 2015; Wilson et al., 2018). In our design, the number markers following the target *gustar*-type verb morpheme were almost always adjacent to one other (e.g., *me molestan los deportes* 'sports annoy me').² Our results thus add to the TE literature by showing that in constructions with adjacent dependencies, there may be no additional benefit in highlighting morphological cues.

RQ2: Effects of Textual Enhancement on Written Elicited Imitation of the Preterite–Imperfect Contrast

For the preterite–imperfect contrast, we demonstrated a significant positive effect of both TE1 and TE2 compared to NC—but not between the two TE conditions. The few studies examining TE effects on the SLA of the preterite–imperfect contrast have yielded mixed findings, with some researchers reporting positive effects of TE on learners' noticing and production of these forms (e.g., Jourdenais et al., 1995; Leeman et al., 1995), while others have not (Cintrón–Valentín et al., 2019; Overstreet, 1998). In the studies showing positive effects, learners in the enhancement condition may have benefited from an added compound enhancement (cf. Han et al., 2008). For example, in Leeman et al. (1995), in addition to receiving TE with combined corrective feedback, learners received enhancement of forms outside of the classroom through take-home assignments that included explicit instructions to focus on both meaning and form while processing the input. Additionally, as

part of the TE in Leeman et al., learners had the opportunity to re-access previously presented text, thus allowing for more permanent visual substance of the textually enhanced forms (see also Jourdenais et al., 1995). In our study, in contrast, learners did not have the opportunity to re-access the previous discourse when viewing an enhanced form—this methodological difference may help to explain the lack of significance between TE1 and TE2 in our results (for further elaboration, see Bardovi-Harlig, 1998).

Regarding studies that have not found positive effects of TE on the SLA of the preterite-imperfect contrast, Overstreet (1998) noted that any lack of TE effects may be due to the greater difficulty of learning how two forms function contrastingly within a specific semantic context as opposed to a single form. Overstreet suggested that TE may be more effective when directed at one grammatical form at a time, instead of when directed at the contrast between the two. Along these lines, Han et al. (2008) noted that there may be a trade-off between focusing learners' attention on the forms enhanced by TE and learners' comprehension of the discourse surrounding the targeted forms. Altogether, the collective findings on the preterite-imperfect contrast open space for more nuanced TE designs that assess the benefits of presenting one form at a time (rather than two), as well as for designs that examine whether learners utilize the opportunity to re-access previous contextual and discourse information.

RQ3: Effects of Textual Enhancement on the Written Elicited Imitation of the Subjunctive in Noun Clauses

The Spanish subjunctive is a relatively complex morphosyntactic structure emerging late in L1 and L2 Spanish acquisition. Contrary to the other target structures in our study, the subjunctive is primarily restricted to subordinate clauses in Spanish (DeKeyser & Prieto

Botana, 2013). In our experiment, the verb in the main clause, acting as a cue to the subordinated subjunctive verb, and the relationship between the two verbs, were made salient through the TE manipulations. We uncovered significant effects of TE1 and TE2 over NC, as well as an added effect of TE2 over TE1. The additional positive effect of TE2 contrasts with what we observed for *gustar*-type verbs and the preterite–imperfect contrast. This suggests that breaking down a sentence’s components and providing learners with structure-specific instructional strategies can improve appropriate mood selection (e.g., Collentine, 2013; Wilson et al., 2018). Our findings furthermore align with previous research that shows that the enhancement of the intervening material between nonadjacent dependencies likewise enhances the salience of the dependencies themselves (cf. Gebhart et al., 2009; Vuong et al., 2015). We demonstrate that this effect is possible through the incorporation of typographical enhancement in captioning.

Responding to the Study’s Hypotheses

For all three target structures, we confirmed our first hypothesis that both TE conditions would render more accurate written reproductions relative to the NC condition. Regarding our second hypothesis that considered the relative effect between the two TE conditions, we found that only the subjunctive received a greater benefit from TE2 compared to TE1. We therefore offer partial confirmation for the second hypothesis: There is a benefit of TE2 on the subjunctive but not on the preterite–imperfect contrast.

Critically, although we demonstrate that TE can help refocus learner attention to nonsalient forms, it does not always follow that this immediate noticing of forms will lead to

their eventual acquisition, or more specifically, to generalized knowledge of these forms through novel recognition and production measures (cf. Roberts & Siyanova–Chanturia, 2013). For example, in research that employs eye-tracking methods to measure learners' visual attention to form, some studies report strong links between noticing and subsequent acquisition (e.g., Indrarathne & Kormos, 2017), whereas others report more nuanced findings (e.g., Godfroid & Uggen, 2013; Winke, 2013). It will thus be imperative for future research to address the potential links between textually enhanced captions, noticing, and L2 learning through longitudinal designs.

Implications for Research on Textual Enhancement

Until recently, the captioning research had focused primarily on its capacity to facilitate vocabulary learning and comprehension, with few studies investigating its potential to support L2 grammar learning (Cintrón–Valentín & García–Amaya, 2021; Cintrón–Valentín et al., 2019;). The results of Cintrón–Valentín et al. (2019) revealed that captioning with TE can aid in the learning of L2 grammar for *gustar*-type verbs and the subjunctive, but not for the *ser–estar* copula contrast or for the preterite–imperfect contrast. Such mixed findings are not uncommon in the TE literature (e.g., Han et al., 2008; Lee & Huang, 2008) and highlight the need to consider TE and captioning effects on a diverse set of grammar structures. Differently from our research, the study by Cintrón–Valentín et al. investigated learner intake of target grammar through a developmental design but did not include a measure of immediate attention through an EI methodology. As mentioned previously, conscious attention to linguistic forms in L2 input is an important precondition for SLA (Schmidt, 2001). The current study thus complements previous research by showing that TE facilitates not only the learning of *gustar*-type verbs and the

subjunctive (as demonstrated in previous research) but also their immediate reproduction.

At the same time, we showed more nuance with respect to the preterite–imperfect contrast: Although we documented an effect of TE on immediate reproduction accuracy, Cintrón–Valentín et al. did not uncover long-term gains through a longitudinal design.

As an additional implication, we contribute to the TE literature by showing that morpheme-enhanced captioning can offer an added positive effect on the reproduction accuracy of some nonadjacent dependencies (in the case of the subjunctive), but not necessarily on adjacent dependencies (in the case of *gustar*-type verbs; cf. Vuong et al., 2015; Wilson et al., 2018). Future research will benefit from further exploring the effect of TE on adjacent versus

nonadjacent dependencies through a greater variety of TE conditions. This would be especially relevant for the preterite–imperfect contrast, for which we found effects of TE1, but not of TE2.

Implications for Research on Teaching L2 Grammar

To encourage a greater variety of approaches to grammar teaching, Larsen–Freeman (2003) called for an increased implementation of ‘grammaring,’ whereby students practice grammar use in circumstances similar to those that they will encounter outside of the classroom. To capture the full effects of grammaring, instructors must consider the specific learning challenges (e.g., complex morphology, meaning, contextual use) posed by different grammar rules and appropriately adjust their classroom practices (Larsen–Freeman, 2009). One clear outcome from our study is that captions with TE constitute a useful tool for L2 instructors.

Critically, the optimal design of the TE manipulation—be it focused on a full lexical entry, the target morpheme, and/or additional sentential cues—should be carefully tailored to the target structure in question. For example, nonadjacent structures such as *if* clauses require high levels of noticing and thus increase overall processing demands (e.g., Rosa & Leow, 2004). In teaching such grammar points, learners will likely benefit from techniques that highlight a balance of syntactic and morphological considerations (see also Uggen, 2012, for a discussion on the learning of complex structures).

As with any instructional method, a single pedagogical technique will not be equally effective for all learners (see Indrarathne & Kormos, 2017; Spada & Tomita, 2010). Factors such as learner proficiency, attitude, motivation, and modality preferences can affect learners' receptiveness to different instructional interventions (Dörnyei, 2005; Muñoz, 2017), leading to considerable between-participant variation (e.g., Larsen–Freeman, 2018). In our results, although all learners demonstrated sensitivity to TE for at least one structure, their individual patterns were not uniform—for example, four learners showed negative effects on four of the six TE comparisons that we conducted. Clearly, much work remains in terms of fine-tuning the quantity and types of enhancement needed for successful grammar acquisition when taking into account diverse groups of learners.

Limitations and Future Directions

This study has its limitations, including (a) the lack of a nonenhanced captioning condition, (b) the inclusion of a single outcome measure limited to the written modality, and (c) the lack of a more direct assessment of prior knowledge. Regarding the first limitation, TE designs that include a straightforward comparison between a NC condition, a nonenhanced captioning condition (absent from this study), and enhanced captioning

conditions would be more equipped to discern whether the use of captioning is the single contributing factor in obtaining positive effects on L2 grammar performance (Leow & Martin, 2017). Given our design, we were unable to differentiate the confounding effects that may have arisen from the written modality of captioning from those stemming from the incorporation of TE in addition to captioning. Regarding the second limitation, Sydorenko (2010) demonstrated that variations in test modality can render differential outcomes on learner performance. We found significant positive effects of written support in a written assessment task but did not include an aural assessment. A next step would involve experiments that consider the relation between input modality and test modality. For the third limitation, studies that probe learners' prior knowledge in more detailed ways (i.e., beyond the use of a global proficiency test) would afford researchers insight into the degree of exposure and prior knowledge needed for successful TE interventions. Such designs would help to determine the extent to which there are correlations between global measures of grammatical proficiency (like our proficiency test) and more local measures of target-structure proficiency. It would likewise be necessary to include a larger sample size for replication purposes.

<A>CONCLUSION

This study examined the role of textually enhanced captions on learners' immediate reproduction of three constructions in L2 Spanish. One methodological innovation was our adaptation of a WEI task to investigate the effect of TE on learners' abilities to reproduce three target grammatical structures: *gustar*-type verbs, the preterite–imperfect contrast, and the subjunctive in noun clauses. Our experimental design focused on the WEI of these three structures, understanding EI as a learner's immediate reproduction of a stimulus

following a targeted intervention. For the three grammar constructions, captions in addition to some form of TE contributed to increased accuracy. This suggests that TE led to greater noticing, or initial detection, of the target features. We showed an additional positive effect of morpheme- and dependency-enhanced TE for the subjunctive in noun clauses. The within-subjects design further allowed us to capture individual performance across the three target features, underscoring the nuance that may arise from the influence of personal variables.

We further laid out a series of implications for L2 researchers and instructors—namely, that (a) incorporating some type of TE leads to increased accuracy in learners' reproduction of target L2 grammar, (b) there are differential effects of TE based on the target structure, (c) the effects of captioning on L2 grammar learning may vary according to factors such as morphosyntactic dependencies, and (d) future research should explore whether there are longitudinal gains on grammar learning (e.g., Indrarathne & Kormos, 2017). Altogether, through increasing the accessibility of multimedia tools in L2 classrooms, TE captions represent a powerful resource for facilitating the learning of myriad grammar structures.

<A>NOTES

¹The breakdown of the 45 items included in the proficiency test was as follows: 14 items tested the preterite–imperfect contrast; 10 tested the subjunctive; 6 tested nonpast verb tenses; 6 tested pronouns; 6 tested gender agreement or propositions; and 3 tested copula verbs. A correlation analysis between additive subscores of the 10 subjunctive and 35 nonsubjunctive items returned a correlation of 0.672. We further found a correlation of 0.618 between the scores of the 14 preterite–imperfect and 31 non-preterite–imperfect

items. Given these outcomes, and to avoid multicollinearity in our model, we included grammar proficiency (per the 45-item test) as the single control variable in the model.

²There were three target sentences (out of the total set of 24) that included the determiner *mucho*.

‘much’ between the verb and subject.

<A>REFERENCES

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, 471–508.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.

Blake, R. J. (2013). *Brave new digital classroom: Technology and foreign language learning*. Washington, DC: Georgetown University Press.

Blyth, C. (2005). From empirical findings to the teaching of aspectual distinctions. In D. Aoun & R. Salaberry (Eds.), *Tense and aspect in Romance languages* (pp. 211–252). Amsterdam: John Benjamins.

Bowden, H. (2016). Assessing second-language oral proficiency for research: The Spanish elicited imitation task. *Studies in Second Language Acquisition*, 38, 647–675.

- Cintrón–Valentín, M. C., & Ellis, N. C. (2015). Exploring the interface: Explicit focus-on-form instruction and learned attentional biases in L2 Latin. *Studies in Second Language Acquisition*, 37, 197–235.
- Cintrón–Valentín, M. C., & Ellis, N. C. (2016). Salience in second language acquisition: Physical form, learner attention, and instructional focus. *Frontiers in Psychology*, 7, 1–21.
- Cintrón–Valentín, M. C., & García–Amaya, L. (2021). Investigating textual enhancement and captions in L2 grammar and vocabulary: An experimental study. *Studies in Second Language Acquisition*. Advance online publication. doi: <https://doi.org/10.1017/S0272263120000492>
- Cintrón–Valentín, M. C., García–Amaya, L., Ellis, N. C. (2019). Captioning and grammar learning in the L2 Spanish classroom. *The Language Learning Journal*, 47, 439–459.
- Collentine, J. (2013). Subjunctive in second language Spanish. In K. L. Geeslin (Ed.), *The handbook of Spanish second language acquisition* (pp. 270–288). Oxford: Wiley Blackwell.
- Comeaux, I., & McDonald, J. L. (2017). Determining the effectiveness of visual input enhancement across multiple linguistic cues. *Language Learning*, 68, 5–45.
- DeKeyser, R. (2005). What makes learning second language grammar difficult? A review of issues. *Language Learning*, 55, 1–25.
- DeKeyser, R., & Prieto Botana, G. (2013). Acquisition of grammar by instructed learners. In

- K. L. Geeslin (Ed.), *The handbook of Spanish second language acquisition* (pp. 449–465). Oxford: Wiley Blackwell.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143–188.
- Ellis, N. C. (2017). Salience in usage-based SLA. In S. Gass, P. Spinner, & J. Behney (Eds.), *Salience in second language acquisition* (pp. 21–40). New York: Routledge.
- Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, 27, 464–491.
- García-Amaya, L. (2012). *Second language fluency and cognition: The study of Spanish secondlanguage development in an overseas immersion program and an at-home foreign language classroom* (Unpublished doctoral dissertation). Indiana University, Bloomington, IN.
- Gass, S. M., Spinner, P., & Behney, J. (Eds.). (2018). *Salience in second language acquisition*. New York: Routledge.
- Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and non-adjacent dependencies among non-linguistic sounds. *Psychonomic Bulletin & Review*, 16, 486–490.

- Ghia, E. (2012). *Subtitling matters: New perspectives on subtitling and foreign language learning*. New York: Peter Lang.
- Godfroid, A., & Uggem, M. (2013). Attention to irregular verbs by beginning learners of German. *Studies in Second Language Acquisition, 35*, 291–322.
- Goldschneider, J. M., & DeKeyser, R. (2001). Explaining the “natural order of L2 morpheme acquisition” in English: A meta-analysis of multiple determinants. *Language Learning, 51*, 1–50.
- Gudmestad, A. (2012). Acquiring a variable structure: An interlanguage analysis of second language mood use in Spanish. *Language Learning, 62*, 373–402.
- Han, Z., Park, E. S., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics, 29*, 597–618.
- Huang, H. C., & Eskey, D. E. (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*, 75–96.
- Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions. *Studies in Second Language Acquisition, 38*, 401–430.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60–99.
- Jourdenais, R., Ota, M., Stauffer, S., Boyson, B., & Doughty, C. (1995). Does textual

enhancement promote noticing? A think-aloud protocol analysis. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 182–209.). Honolulu, HI: University of Hawai'i Press.

LaBrozzi, R. M. (2016). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research, 20*, 75–91.

Larsen–Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Boston, MA: Thomson/Heinle.

Larsen–Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*, 579–589.

Larsen–Freeman, D. (2018). Looking ahead: Future directions in, and future research into, second language acquisition. *Foreign Language Annals, 51*, 55–72.

Lee, J. F., & Malovrh, P. A. (2009). Linguistic and non-linguistic factors affecting OVS processing of accusative and dative case pronouns by advanced L2 learners of Spanish. In J. Collentine, M. García, B. Lafford, & F. Marcos Marín (Eds.), *Selected proceedings of the 11th Hispanic Linguistics Symposium* (pp. 105–116). Somerville, MA: Cascadilla Proceedings Project.

Lee, M., & Révész, A. (2018). Promoting grammatical development through textually enhanced captions: An eye-tracking study. *Modern Language Journal, 102*, 557–577.

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*, 307–331.

- Leeman, J., Arteagoitia, I., Fridman, B., & Doughty, C. (1995). Integrating attention to form with meaning: Focus on form in Spanish content-based instruction. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 217–258). Honolulu, HI: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Leow, R. P., & Martin, A. (2017). Enhancing the input to promote salience of the L2: A critical overview. In S. Gass, P. Spinner, & J. Behney (Eds.), *Salience in second language acquisition* (pp. 167–186). New York: Routledge.
- Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*, 673–680.
- Liskin-Gasparro, J. (2000). The use of tense–aspect morphology in Spanish oral narratives: Exploring the perceptions of advanced learners. *Hispania*, *83*, 830–844.
- Markham, P. L. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals*, *32*, 321–328.
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). Open Sesame: An open source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*, 314–324.
- Mayer, R.E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction*, *13*, 125–139.
- 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.
- Muñoz, C. (2017). The role of age and proficiency in subtitle reading. An eye-tracking study.

System, 67, 77–86.

- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I: Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127–162.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417–528.
- Overstreet, M. (1998). Text enhancement and content familiarity: The focus of learner attention. *Spanish Applied Linguistics, 2*, 229–258.
- Plass, J. L., & Jones, L. C. (2005). Multimedia learning in second language acquisition. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 467–488). New York: Cambridge University Press.
- Plonsky, L., & Oswald, F. L. (2014). How big is ‘big’? Interpreting effect sizes in L2 research. *Language Learning, 64*, 878–912.
- Roberts, L., & Siyanova–Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition, 35*, 213–235.
- Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21*, 45–105.
- Rosa, E., & Leow, R. (2004). Computerized task-based exposure, explicitness, type of feedback, and Spanish L2 development. *Modern Language Journal, 88*, 192–216.
- RStudio Team (2015). *RStudio: Integrated development for R*. RStudio, Inc., Boston, MA.

- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
- Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, *Proceedings of CLaSIC 2010, Singapore, December 2–4* (pp. 721–737). Singapore: National University of Singapore, Centre for Language Studies.
- Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15, 165–179.
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60, 263–308.
- Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology*, 14, 50–73.
- Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183–203.
- Tracy–Ventura, N., McManus, K., Ortega, L., & Norris, J. M. (2014). Repeat as much as you can: Elicited imitation as a measure of global proficiency in L2 French. In P. Leclercq, H. Hilton, & A. Edmonds (Eds.), *Proficiency assessment issues in SLA research: Measures and practices* (pp. 143–166). Bristol, UK: Multilingual Matters.
- Tragant Mestres, E., & Pellicer–Sánchez, A. (2019). Young EFL learners' processing of multimodal input: Examining learners' eye movements. *System*, 80, 212–232.

- Uggen, M. (2012). Reinvestigating the noticing function of output. *Language Learning*, 62, 506–540.
- Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching*, 43, 1–37.
- VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood, NJ: Ablex.
- Vázquez Rozas, V. (2006). *Gustar*-type verbs. In J. C. Clements & J. Yoon (Eds.), *Functional approaches to Spanish syntax: Lexical semantics, discourse and transitivity* (pp. 80–114). New York: Palgrave MacMillan.
- Vinther T. (2002). Elicited imitation: A brief overview. *International Journal of Applied Linguistics*, 12, 54–73.
- Vuong, L., Mayer, A. S., & Christiansen, M. H. (2015). Concurrent statistical learning of adjacent and nonadjacent dependencies. *Language Learning*, 66, 8–30.
- Wilson, B., Spierings, M., Ravignani, A., Mueller, J. L., Mintz, T. H., Wijnen, F., van der Kant, A., Smith, K., & Rey, A. (2018). Non-adjacent dependency learning in humans and other animals. *Topics in Cognitive Science*, 12, 843–858.
- Winke, P. M. (2013). The effects of input enhancement on grammar learning and comprehension. *Studies in Second Language Acquisition*, 35, 323–352.
- Winke, P. M., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14, 65–86.

Winke, P. M., Gass, S., & Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *Modern Language Journal, 97*, 254–275.

Yan, X., Maeda, Y., Lv, J., & Ginther, A. (2016). Elicited imitation as a measure of second language proficiency: A narrative review and meta-analysis. *Language Teaching, 33*, 497–528.

Author Manuscript