The Effects of Form Focused Instruction and Captioning on Second Language Development

by

Myrna C. Cintrón Valentín

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in the University of Michigan
2019

Doctoral Committee:

Professor Nick Ellis, Chair
Professor Julie Boland
Assistant Professor Jon Brennan
Assistant Professor Lorenzo García-Amaya

Myrna C. Cintrón Valentín

mcintron@umich.edu

ORCID iD: 0000-0002-7560-7566

DEDICATION

A mis queridos padres, por su apoyo incondicional

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor, Nick Ellis, under whose direction this work has been carried out, for his invaluable support and advice during the elaboration of this dissertation and throughout my journey at the University of Michigan.

Next, I would like to thank all of the members of my dissertation committee, Julie Boland, Lorenzo García-Amaya, and Jon Brennan, for their service, scientific insights, and valuable feedback given throughout the dissertation process. Each one, from their area of study, has greatly contributed to the preparation of this research and my academic training.

A special thanks to Nicholas Henriksen for his generous feedback during the development of the studies included here and for contributing to both my academic and professional training. Thank you as well to Brady West and Micha Fisher for their valuable advice in the statistical analyses included in this dissertation.

I would also like to thank all of my past and current Research Assistants for their hard work and dedication to this project– this research would not be possible without them.

Thank you to my CCN family for their unwavering support, encouragement, and friendship throughout my time in Michigan. I could not have made it in one piece without you!

To my family– Humberto Cintrón Aybar, Rosalind Valentín Álvarez, and Carmen Cintrón Valentín– and my partner, Luis Velázquez Irizarry, without your love and support throughout this process none of this would be possible. Gracias infinitas.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1256260.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

An essential component in the L2 acquisition process is conscious attention to form in the input (Schmidt, 2001). Given that some linguistic forms are inherently less noticeable (e.g., grammatical forms), a central question in the field of Second Language Acquisition (SLA) is how learner attention can be drawn to such linguistic features in order to promote learning. Traditionally, attention-getting strategies such as Textual Enhancement (TE) or Explicit Grammar Instruction (EGI) – two types of Form Focused Instruction (FFI) – have been used to counteract these low salience effects. The usage of these techniques, however, has for the most part been limited to the written modality with few studies investigating the role of multimodal input (i.e., aural, visual and pictorial input) in facilitating grammar acquisition. One promising multimodal technique which has been the focus of much recent research is that of captioned video. Extensive research from the last three decades has demonstrated its effectiveness in L2 comprehension and vocabulary learning (Vanderplank, 2010). However, little attention has been paid to its potential in supporting grammar learning, a challenging area of L2 acquisition. The studies presented in this dissertation aim to extend previous research on captioning and L2 acquisition by targeting grammar learning. They additionally build upon existing research by exploring how FFI techniques such as EGI in combination with captioned video, and salience-raising manipulations through TE within the caption line might aid in facilitating grammar development.

Studies 1 and 3 of this dissertation explore the role of FFI + captioned media in the L2 Spanish classroom through two separate random-allocation field experiments. Altogether, the findings of

these studies confirm the effectiveness of captioning on vocabulary, illustrate the extra difficulties of grammar, and help inform which types of constructions might be assisted by captioning. Study 2 of this dissertation assesses effects of different designs of TE video captions on learners' immediate uptake of grammatical constructions in L2 Spanish. The findings of Study 2 offer key methodological insights for fine-tuning the amount of enhancement that might be required for successful learner uptake through TE.

CHAPTER 1. General Introduction

Second language (L2) acquisition is a complex sociocognitive process, in many ways different from that of first language (L1) acquisition given the more abundant hurdles involved in its development. To begin with, learners come to the task of acquiring their L2 following that of their L1–thus their history of language experience inevitably influences their L2 learning outcomes. At the same time, the process of acquiring a L1 is typically accomplished during the first few years of a child's life, whereas L2 development can begin at any stage following L1 acquisition and continue throughout an individual's lifespan. Additionally, the learning contexts to which L1 and L2 learners are exposed to often differ – whereas L1 learners are generally exposed to multiple sources of input and in varying contexts, L2 learners' exposure to, and usage of the target language, is typically limited to the classroom environment. Further, as the Second Language Acquisition (SLA) literature demonstrates, even learners who have acquired their L2 through input-rich naturalistic environments often struggle in reaching high levels of proficiency when it comes to learning the grammar of their L2. What factors might then contribute to this generalized difficulty in L2 acquisition? And how might learners overcome these hurdles?

The present dissertation investigates one possible avenue for advancing the learning of a second language. Specifically, by integrating theoretical principles from the SLA grammar learning literature with multimedia learning methods, I present a series of studies investigating how captioned media in combination with Form Focused Instruction (FFI) techniques might facilitate the acquisition of grammatical forms in the input.

In the sections that follow, I will summarize several central concepts and theories within SLA which attempt to explain how the process of L2 acquisition takes place. I will begin by summarizing how input has been conceptualized in the field of SLA. I will additionally discuss several factors which have been implicated in mediating L2 learners' processing of input, namely those of attention, perceptual salience, FFI, and modality of input presentation. Finally, I will provide an overview of the Chapters presented in this dissertation.

1. The Need for Input in L2 Acquisition

Input can be defined as the language – aural, written or visual–a learner is exposed to, and thus "constitutes the data that learners have to work with to construct their interlanguage[1]" (Ellis, 2015). Several foundational theories within SLA recognize the importance of input in L2 acquisition, although they differ in their understanding of how the processes of acquisition takes place (Long, 1996; Krashen, 1985; Swain, 1995). The *Input Hypothesis* (Krashen, 1985), for instance, claims that acquisition takes place through an unconscious automatic process based solely on learners' exposure to comprehensible input. Comprehensible input is understood as simplified "input that contains language slightly beyond the current level of the learner's internalized language" (Gass & Mackey, 2014, p. 26).[2] L2 acquisition through this account is assumed to be entirely input-driven paralleling the process of L1 acquisition in which the 'building up' of the learners' L1 grammar is largely influenced by naturalistic exposure to their caregiver's speech in meaningful contexts (Ortega, 2009).

[1] Interlanguage refers to the language system that each learner construct at any given point in development (Selinker, 1972)

[2] One example of a comprehensible input approach is the usage of graded readers –books written especially for foreign language students to facilitate the acquisition process (see for instance, Rodrigo, Krashen & Gibbons, 2004).

However, as extensive research in SLA has demonstrated (e.g., Long, 1990; Schmidt, 1994, 2001), comprehensible input alone may not always suffice for the process of L2 acquisition to take place. Swain (1985), for instance, investigated L2 French immersive programs in Canada, finding that despite exposure to abundant L2 input, learners still did not achieve high levels of grammatical and sociolinguistic competence. In her view, the missing component in these immersive contexts was the provision of more abundant and meaningful opportunities for target language usage by way of pushed output (i.e., output where learners 'push' themselves to be more comprehensible during the process of communication). Through the *Output Hypothesis*, Swain (1985) proposed that in addition to input, "producing the target language may be the trigger that forces the learner ***to pay attention*** to the means of expression needed in order to successfully convey his or her own intended meaning" (p. 249, emphasis added). The *Interaction Hypothesis* (Long, 1983, 1996) similarly emphasizes the role of both input and output for successful acquisition. Long (1983) specifically proposed that the most optimal form of comprehensible input is that which has been interactionally modified or adjusted through a process of negotiation of meaning between the interlocutors in order to convey a more comprehensible message. Long (1996) asserts that "negotiation of meaning, and especially negotiation work that triggers interactional adjustments by the native speaker or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, ***particularly selective attention***, and output in productive ways" (pp. 451-452, emphasis added).

Although not explicitly recognized in their earlier formulations, both the *Output and Interaction hypotheses* highlight one additional component of L2 acquisition: the importance of

noticing and ***conscious attention to linguistic forms*** in the input. In their current versions, there is a role for both components during the process of communication and negotiation of meaning. The notion of ***noticing*** and ***conscious attention to linguistic forms*** in the input as an essential component in the learning process, however, is more clearly outlined through Schmidt's *Noticing Hypothesis* (1994, 2001).

2. A Role for Noticing and Conscious Attention to Input

The *Noticing Hypothesis* is viewed as one of the most influential theoretical accounts in SLA given its contribution to the understanding of the role of attention in the L2 learning process. The main theoretical premise behind Schmidt's (1994; 2001) *Noticing Hypothesis* is that conscious attention to linguistic forms (e.g., sounds, words, grammar) in the input is an important precondition to learning. Specifically, Schmidt stresses learners must notice linguistic forms in the input in order for them to become intake for learning. Intake can be defined as the subset of input that has been processed in some way by the learner […] [and that] is created when learners make form-meaning connections from the input" (Wong, 2005, p. 119). Attention to linguistic forms, as framed by the Noticing Hypothesis, does not need to be intentional (i.e., deliberate or goal-directed) – it can also occur incidentally (e.g., when learning new vocabulary while reading for comprehension rather than acquisition). Nonetheless, as Schmidt (2001) asserts, deliberate attention to form may be necessary in some cases, for example, in the acquisition of morphological and syntactic features which may not be immediately noticeable in the input. This observation in now commonplace in the SLA literature (e.g., Bardovi-Harlig, 1992; Clahsen & Felser, 2006; VanPatten, 1996), with a large body of work investigating under what conditions learners may notice specific linguistic forms in the input, and whether more noticing leads to

greater learning gains (e.g., Cintrón-Valentín & Ellis, 2015; Godfroid, Boers & Housen, 2013; Indrahane & Kormos, 2017; Leow, 2001; Simard & Foucambert, 2013; Winke, 2013).

Noticing facilitates learners' ability to modify pre-existing linguistic knowledge by consciously comparing the differences between target-like input and their own un-target-like output. It is through this process that learners may begin to encode and eventually accommodate the linguistic form into their long-term memory. Thus, attention is viewed as a necessary condition to learning, particularly at early stages of the L2 learning process.

3. Defining the Problem space: The Shortcomings of L2 Acquisition

Perhaps, one of the most representative studies illustrating the generalized difficulty of acquiring grammatical forms in L2 development is that of Klein and Perdue (1992). For a period of two and a half years, the researchers followed 40 L2 learners, who varied in their native and target languages in order to assess their naturalistic learning of the languages in question (i.e., English, German, Dutch, French, and Swedish). Surprisingly, rather than finding marked differences in their L2 development, they found a similar acquisitional pattern in their corresponding interlanguages. The majority of the learners developed what was coined as the Basic Variety, an interlanguage described as a simple learner language characterized by a high use of lexical items, little to no use of closed-class items, no use of functional inflections and a greater reliance on lexis and pragmatic devices in their expression of temporality. Similar findings have been revealed in a number of studies investigating both naturalistic and classroom instruction (e.g., Bardovi-Harlig, 1992; Ellis, 1989). These overall findings demonstrate that not all L2 learners go beyond pragmatic and lexical stages of language usage in their L2 development and is a reflection of the difficulties encountered by adult L2 learners in general – typically, they may

learn more words, but grammatical abilities tend to plateau and do not progress into high levels of attainment. Although this interlanguage is adequate for everyday purposes, the Basic Variety falls short of native-like competence. One concrete example of this Basic Variety phenomenon can be found in Schmidt's (1983; 1984) case study of Wes, a Japanese learner of English whose L2 development was recorded over the course of five years. Wes was described as very fluent with high levels of strategic competence, but low levels of grammatical accuracy. Schmidt (1984), specifically notes:

> If language is seen as a medium of communication, as a tool for initiating, maintaining and regulating relationships and carrying on the business of life, then W [referring to Wes] has been a successful language learner... If language acquisition is taken to mean (as it usually is) the acquisition of grammatical structures, then the acquisition approach may be working, but very slowly. Using 90% correct in obligatory contexts as the criterion for acquisition, none of the grammatical morphemes counted has changed from unacquired to acquired status over a five-year period" (p. 5).

Why do L2 learners favor lexical and pragmatic means over grammatical forms in the input? What factors might mediate their attentional focus to these forms during input processing and their subsequent learning? The SLA literature provides various accounts on L2 learners' difficulty in acquiring various aspects of grammar, including critical periods for language acquisition (e.g., Johnson & Newport, 1989; Hartshorne et al., 2018), input processing differences between native speakers and L2 learners (VanPatten, 1996; 2003), individual differences (e.g. personality differences, cognitive differences, as well as differences in learners'

personality traits; Dörnyei, 2005), as well as the linguistic features of the target grammar constructions themselves, such as frequency (e.g., Ellis, 2002; 2006), and perceptual salience (e.g., Ellis, 2017; Gass, Spinner & Behney, 2017; Goldschneider & DeKeyser, 2001; Larsen-Freeman, 1976). The focus of this dissertation will be on the latter, specifically that of perceptual salience, given its more prominent role in mediating learner attention to linguistic forms in the input (e.g., Gass et al., 2017).

4. Salience in L2 Acquisition

SLA research demonstrates that regardless of the vast availability of grammatical forms in the input, L2 learners quite often ignore these forms during input processing, and focus more upon open-class words, such as nouns, verbs, adjectives and adverbs (e.g., Bardovi-Harlig, 1992; Clahsen & Felser, 2006; Schmidt, 2001). One reason why grammatical forms might prove impervious to L2 learners is due to their low perceptual salience in the input. Perceptual salience refers to the intrinsic qualities of a linguistic form or structure, e.g., amount of phonetic substance, stress level, usual serial position, etc. (see for instance, Brown, 1973; Goldschneider & DeKeyser, 2001). Many grammatical form-function relationships in English, such as grammatical particles and inflections (e.g., the third person singular *-s* or past tense *-ed*) tend to be short and low in stress, with the result that these cues are difficult to perceive, 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 & DeKeyser, 2001). Encouraging learners to attend to these linguistic forms, in many cases, will thus require the provision of salience-raising techniques designed to increase their learnability.

5. Enhancing Attention to Non-Salient Forms

5.1. The role of Form Focused Instruction

FFI is a term describing the wide range of instructional activities that look to draw learners' attention to linguistic forms in the input that might otherwise be ignored. One example, Processing Instruction (VanPatten, 1996), aims to alter learners' default processing strategies to change the ways in which they attend to input data and, thus, to maximize the amount of intake of data in L2 acquisition. Options for FFI (see for instance, Norris & Ortega, 2000) vary in a continuum ranging from those that are more explicit to those that are more implicit. DeKeyser (1995) argues that explicit instruction requires that there be "some sort of rule…being thought about during the learning process" (p.380,), and is thus aimed at encouraging metalinguistic understanding of specific target structures. This type of instruction can be deductive, when learners are presented with a particular rule, or inductive, when they are asked to attend to a particular set of forms with the purpose of inferring the rules on their own. Conversely, in implicit instruction, learners are neither given a rule, nor asked to infer rules from the input learners, and are thus expected to infer rules without awareness. In this case, they are neither given a rule, nor asked to infer rules from the input (R. Ellis, 2012; Norris & Ortega, 2000). Below, I summarize the two FFI techniques that will be the focus of this dissertation.

5.1.1. Grammar Instruction

One method that has been widely used in SLA research is that of grammar instruction. Terrell (1991) defines explicit grammar instruction (EGI) as "the use of instructional strategies to draw the students' attention to, or focus on, form and/or structure" (p. 53), with instruction targeted at increasing the salience of inflections and other commonly ignored features by, first, pointing

them out and explaining their structure and, second, providing meaningful input that contains many instances of the same grammatical meaning-form relationship. Benefits for this type of instruction were found in a meta-analysis on the effectiveness of FFI by Norris and Ortega (2000), where the trend for explicit treatments suggested that instructional conditions, which involved a focus on the rules underlying specific L2 structures led to greater advantages in learning than those that did not include such a focus.

5.1.2. Textual Enhancement (TE)

One other well-known FFI technique is that of Input Enhancement (IE), described as an unobtrusive method aimed at enhancing learners' awareness of non-salient forms in the input (Doughty & Williams, 1998; Sharwood Smith, 1993). There are a number of IE techniques used for enhancing both visual and oral input, including but not limited to gestures, intonation and textual manipulations. In Textual enhancement (TE), visual manipulations, such as color-coding, boldfacing, and underlining, are typically used to enhance forms in written input, and therefore facilitate learners' further processing of these cues (Sharwood Smith, 1993).

Research in this area has yielded conflicting findings regarding its effectiveness, some demonstrating that TE is successful in drawing learners' attention to the target forms (e.g. locative suffixes, preterit and imperfect verb forms, relativization, and passive constructions) (Alanen, 1995; Cho, 2010; Izumi, 2002; Jourdenais, 1995; Lee, 2007; Winke, 2013), and in learners' subsequent learning of these forms (Jourdenais, 1995; Lee, 2007; Shook, 1994), as measured by a variety of tests, including recognition tasks, and recall and grammaticality judgment tasks. However, other studies, have found no effect of enhancement on learning (Izumi, 2002; Leow, 1997, 2001; Leow, Egi, Nuevo, & Tsai, 2003; Overstreet, 1998; Wong,

2003). These discrepancies may be explained by methodological differences across the TE literature, including differences in learners' target and native languages, the type and amount of target forms in each study, and how the TE manipulations are realized (Han, Park, & Combs, 2008; Lee & Huang, 2008). These issues will be further discussed in Chapter 3 of this dissertation, where given that they can be successful, I try to identify how to optimize FFI for L2 learning of Spanish grammar.

6. Salience and Modality of Input Presentation

Spoken and written language are very different media, with spoken language being fleeting while written language provides more permanent visual substance on the page, allowing the reader to attend to linguistic forms at their discretion. Attention to language form may therefore pose different challenges in written and spoken modalities, and acquisition is usually superior from visual input (Cintrón-Valentín & Ellis, 2016; Morgan-Short et al, 2018; VanPatten, 1990; Wong, 2001; Vidal, 2011). Cintrón-Valentín and Ellis (2016), for instance, found that across different types of instructional treatments such as IE and EGI, L1 English learners of Latin were more able to acquire temporal reference through the written rather than aural modality. Similarly, Vidal (2011), who targeted L1 Spanish learners of English, found an advantage in gains associated with vocabulary acquisition for students presented with an academic text versus a video-taped lecture containing the same content. Thus, modality can differentially affect the salience of forms and their input processing: written language can make grammatical forms more salient and more easily processed.

Research on modality effects in L2 acquisition has also found that providing learners with bimodal/multimodal input, i.e., enriching the aural with written and/or visual cues, can lead to significant advantages over aural input alone in both vocabulary acquisition (e.g., Bird & Williams, 2002; Jones and Plass, 2002; Montero-Pérez et al., 2013 Webb & Nation, 2017), and grammar development (e.g., Cintrón-Valentín, 2016; Lee & Révész , 2018). For example, in a follow-up study to Cintrón-Valentín and Ellis (2016), Cintrón-Valentín (2016) found robust learning effects for Visual and Aural-Visual modalities against the Aural only group for both TE and EGI groups in learners' acquisition of verb-tense morphology. Further, Jones and Plass (2002) presented learners with an aural passage through four conditions: no annotations, written or pictorial annotations alone, or both types of annotation. Learners in the dual-annotation condition outperformed all other groups, whereas those who received either written or pictorial annotations performed better than the no-annotations group but did not significantly differ from each other. Webb and Nation (2017) discuss how the use of elaboration techniques, designed to enrich a learner's knowledge of a word "by encountering more aspects of its form, meaning, and use", such as the inclusion of pictures in addition to written text, can in many instances ,"provide a memorable image of the meaning and context of a word"(p.73), and thus facilitate acquisition.

One particular multimodal/multimedia resource that has been the focus of much recent research within SLA is that of captioned video (i.e., video including subtitles where the text is presented in the same language as the audio) given its demonstrated benefits in facilitating L2 comprehension and vocabulary acquisition (e.g., Vanderplank, 2010; Montero-Perez, et al., 2013). So far, however, little is known about the role of captioning in facilitating L2 grammar acquisition. Captioned video holds special promise for grammar development, given its potential role in mediating learners' attention to specific word-forms in the input (Montero-Perez, 2014;

Vanderplank, 2016; Winke, Gass, & Sydorenko, 2010). Additionally, given the increased reliance on multimedia materials in L2 teaching and learning (Blake, 2013; Plass & Jones, 2005), it is of growing interest to investigate how traditional instructional techniques such as FFI can be effectively integrated with such multimedia approaches to language learning.

7. Current research and overview of the present dissertation

The studies presented in this dissertation aim to extend previous research on captioning and L2 acquisition by targeting grammar learning. They additionally build upon existing research by exploring how FFI techniques such as EGI in combination with captioned video, and salience-raising manipulations through TE within the caption line might aid in facilitating grammar development.

Chapter 2 of this dissertation presents a first study aimed at exploring the role of FFI + captioned media in the L2 Spanish classroom. The study was integrated into a one-semester university L2 Spanish grammar course in a random-allocation field experiment. Through four data-collection sessions, we targeted: the preterite-imperfect contrast, *gustar*-type verbs, copula verbs, and the subjunctive. In each session, participants saw a short grammar lesson before an animated video. The animation video content was constant, but the caption format varied, such that participants were either presented with captions which included textually enhanced target vocabulary, or textually enhanced target grammar, or no captioning was provided. Participants were then tested on their recognition and production of the target items. Results showed clear effects of captioning on vocabulary recognition and production. For some grammar structures, there were also positive effects of captioning on production, whereas for other grammar structures no such effect was uncovered. Altogether, the findings of Study 1 confirm the effectiveness of captioning on vocabulary, illustrate the extra difficulties of grammar, and help

inform which types of construction might be assisted by captioning. Chapter 2 was submitted to The Language Learning Journal and received a 'revise and resubmit'. The revised version of this manuscript has just been resubmitted.

Chapter 3 of this dissertation assesses effects of different designs of TE video captions on learners' immediate uptake of grammatical constructions in L2 Spanish. Through a within-subjects design, L2 Spanish learners saw three animated videos focusing on: *gustar*-type verbs, the preterite-imperfect contrast, and the subjunctive. Each video included three different TE manipulations: control sentences with no captions (NC); TE1-type sentences with target verbs highlighted in their entirety, or TE2-type sentences where only the critical morphological and grammatical cues, and their relations were highlighted. There were clear and significant effects of TE over NC on grammar uptake, which differed by structure and TE-type. Overall, the findings suggest that TE can be improved if it goes beyond mere highlighting of structures to additionally show the grammatical relations between their parts. The findings of Study 2 offer key methodological insights for fine-tuning the amount of enhancement that might be required for successful learner uptake through TE. Chapter 3 has been submitted to Applied Linguistics. We await a response from the journal editors.

Chapter 4 of this dissertation presents a direct follow-up to Study 1 in an effort to better understand the role of captioned media on L2 grammar development. Study 3 specifically addresses several limitations presented in Study 1, namely, the lack of a pretest, and the inclusion of a comparison group which did not receive explicit instruction. Similar to Study 1, Study 3 was integrated into a one-semester university L2 Spanish grammar course. Participants were presented with an initial grammar lesson on the target structure and an animated video, which varied in terms of the captioning format, such that participants were either presented with

captions which included textually enhanced target vocabulary (Lesson + Salience on Vocabulary), or textually enhanced target grammar (Lesson + Salience on Grammar; No Lesson + Salience on Grammar), or no captioning was provided (Lesson + Control). The No Lesson + SG group was not presented with the initial grammar lesson in order to investigate how facilitative the textually enhanced captions would be in the absence of explicit instruction. Following the video presentation, participants were tested on their production of the target vocabulary and grammar items. The vocabulary results from Study 1 were replicated in Study 2. The results additionally showed partial confirmation that there is an effect of captioning on grammar acquisition, although it varied by structure. Taken together, the findings of Study 3 confirm the effectiveness of captioning on vocabulary and show that the acquisition of some grammar structures is more easily facilitated by captioning and TE than others. Chapter 4 has just been finalized as a paper and I intend to submit it to a leading Applied Linguistics Journal within the next two months.

This body of work provides a detailed picture on how FFI + captioned media could be useful in second language development and education. The theoretical and pedagogical implications of this work, as well as future directions, are discussed in **Chapter 5.**

8. References

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259-302). Honolulu: University of Hawaii Press.

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*(4), 299-320.

Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics, 23*(4), 509-533.

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

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Cho, M. Y. (2010). The effects of input enhancement and written recall on noticing and acquisition. *Innovation in Language Learning and Teaching, 4*(1), 71-87.

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*(2), 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, 1284.*

Cintrón-Valentín, M. C. (2016). Salience in Second Language Acquisition. Poster presented at Psychonomics, Boston.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*(1), 3-42.

DeKeyser, R. M. (2005). What makes learning second language grammar difficult? A review of issues. *Language Learning, 55*(1), 1-25.

Dörnyei, Z. (2005). *The Psychology of the language learner: Individual differences in second language acquisition. Mahwah*, NJ: Lawrence Erlbaum.

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. New York, NY: Cambridge University Press.

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*(2), 164–94.

Ellis, N. C. (2017). Salience in usage-based SLA. In S. Gass, P. Spinner, & J. Behney (Eds), *Salience in second language acquisition (*pp. 21-39*)*. New York: Routledge.

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, R. (1989). Classroom learning styles and their effect on second language acquisition: A study of two learners. *System*, 17, 249-262.

Ellis, R. (2012). Form-focused instruction and second language learning. *Language Teaching Research and Pedagogy* (pp. 271-306). US: Wiley-Blackwell.

Ellis, R. (2015). *Understanding Second Language Acquisition* (2nd ed.). Oxford University Press.

Gass, S. M., Spinner, P., & Behney, J. (Eds.). (2018). *Salience in second language acquisition*. New York: Routledge.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye- tracking. *Studies in Second Language Acquisition, 35,* 483-517.

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), 1-50.

Han, Z., Park, E. S., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics, 29*(4), 597-618.

Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions. *Studies in Second Language Acquisition, 38*(3), 401-430.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*(4), 541-577.

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.

Jones, L., & Plass, J. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal, 86*(4), 546-561.

Jourdenais, R. M., Ota, S. Stauffer, B. Boyson, and 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: University of Hawaii Press.

Klein, W., & Perdue C. (1992). Utterance structure. *Developing grammars again*. John Benjamins: Amsterdam/Philadelphia.

Krashen, S. (1985). *The input hypothesis*. London: Longman.

Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning, 26,* 125-134.

Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning, 57*(1), 87.

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*(3), 307-331.

Lee, M., & Révész, A. (2018). Promoting grammatical development through textually enhanced captions: An eye-tracking study. *The Modern Language Journal, 102*(3), 557-577.

Leow, R. P. (1997). The effects of input enhancement and text length on adult L2 readers' comprehension and intake in second language acquisition. *Applied Language Learning, 8,* 151-182.

Leow, R.P. (2001). Do learners notice enhanced forms while interacting with the L2? An online and offline study of the role of written input enhancement in L2 reading. *Hispania, 84,* 496-509.

Leow, R. P., Egi, T., Nuevo, A. M., & Tsai, Y. C. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning, 13,* 93-108.

Long, M. H. (1983). Linguistic and conversational adjustments to nonnative speakers. *Studies in Second Language Acquisition, 5,* 177-193.

Long, M. H. (1990). The least a second language acquisition theory needs to explain. *TESOL*

Quarterly, *24*, 649-666.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York: Academic Press.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41*(3), 720-739.

Morgan-Short, K., Marsden, E., Heil, J., Issa, B., Leow, R. P., Mikhaylova, A., Mikołajczak, S., Moreno, N., Slabakova, R., & Szudarski, P. (2018). Multi-site replication in SLA research: Attention to form during listening and reading comprehension in L2 Spanish. *Language Learning,68*(2), 392-437.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417.

Ortega, L. (2009). *Understanding Second Language Acquisition*. New York: Routledge.

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.

Schmidt, R. (1983). Interaction, acculturation and the acquisition of communicative competence: A case study of an adult. In N Wolfson, & E. Judd (Eds.), *Sociolinguistics and Language Acquisition* (pp.137-174). Rowley, MA Newbury House

Schmidt, R. (1984). The strengths and limitations of acquisition: A case study of an untutored language learner. Language, Learning, and Communication, 3, 1-16.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. In J.H. Hulstijn, & R. Schmidt (Eds.), *Consciousness in Second Language Learning* (pp. 11-26). AILA Review.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*, 165-179.

Shook, D. J. (1994). FL/L2 reading, grammatical information, and the input to intake phenomenon. *Applied Language Learning, 5*, 57-93.

Simard, D., & Foucambert, D. (2013). Observing noticing while reading in L2. In J. M. Bergsleithner, S.N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 207–226). Honolulu, HI: National Foreign Language Resource Center, University of Hawai'i at Mānoa.

Swain, M. K. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass, & C. G. Madden (Eds.), *Input in Second Language Acquisition* (pp. 236-244). Rowley, MA: Newbury House.

Terrell, T. (1991). The role of grammar instruction in a communicative approach. *Modern Language Journal, 75*, 52-63.

Vanderplank, R. (2016). 'Effects of' and 'effects with' captions. How exactly does watching a TV programme with same-language subtitles make a difference to language learners? *Language Teaching, 49*(2), 235-250.

VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 12*(3), 287-301.

VanPatten, B., & Oikkenon, S. (1996). Explanation vs. structured input in processing instruction. *Studies in Second Language Acquisition, 18*(4), 495-510.

VanPatten, B. (2003). *From input to output: A teacher's guide to second language acquisition*. New York: McGraw-Hill.

VanPatten, B., Williams, J. (2014). Introduction: The nature of theories. In B. VanPatten, & J. Williams (Eds.), *Theories in Second Language Acquisition: An introduction* (pp. 61-75). New York: Routledge.

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

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford: Oxford University Press.

Winke, P. M. (2013). The effects of input enhancement on grammar learning and comprehension. *Studies in Second Language Acquisition, 35*(2), 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*(1), 65-86.

Wong, W. (2001). Modality and attention to meaning and form in the input. *Studies in Second Language Acquisition, 23*(3), 345-368.

Wong, W. (2003). Textual enhancement and simplified input effects on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning, 13*(2), 109-132.

Wong, W. (2005). *Input enhancement: From Theory and Research to the classroom*. New York: McGraw-Hill.

CHAPTER 2. To What Extent Can Captioning Facilitate Second Language Acquisition?

1. Introduction

There is increasing interest in the use of captioned videos as a means to promote comprehension and vocabulary learning (Vanderplank, 2010; Montero-Perez, Van den Noortgate & Desmet, 2013). 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 programs, and the internet in L2 classrooms (Blake, 2013; Plass & Jones, 2006; 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 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 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, realizing 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 & Eskey, 1999; Neuman & Koskinen, 1992). These two areas remain the focus of current research (e.g., Muñoz, 2017; Sydorenko, 2010; Winke, Gass, & Sydorenko, 2010; Winke, Sydorenko & 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 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 & Lafayette, 1994, p.70). Likewise, Garza (1991, p. 246) argues that subtitles might help learners "build their aural comprehension in relation to their reading comprehension" because working with this type of format will help them recognize the aural cue of a captioned expression the next time they encounter it. 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 & Behney, 2017; Schmidt, 2001; Tomlin & 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, p.30). Vanderplank's (2016) model of language acquisition through captioned media similarly emphasizes 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 salience[3] as it relates to the perceptual distinctiveness of a linguistic cue in the input has received increasing interest in recent years (Gass, Spinner & Behney, 2017; Ellis, 2006, 2017; Wulff & Ellis, 2018): "salient items or features are attended,

[3] 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 Gass et al., 2017; Ellis, 2017).

are more likely to be perceived, and are more likely to enter into subsequent cognitive processing and learning" (Ellis, 2017, p.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 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.

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 & 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 & DeKeyser, 2001). The use of salience-inducing Input Enhancement manipulations (Doughty & 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, & Combs, 2008; Lee & Huang, 2008; Leow & 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.

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 they 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 media 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.

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:

(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.

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 can 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 words were made salient; and a Captions + TE Grammar group, designed to raise the learner's awareness and attention to grammatical cues. Motivated in part by the findings of Lee and Huang (2008), we targeted four grammatical topics: (1) preterite and imperfect forms, (2) *ser* and *estar* (i.e., copula verbs), (3) *gustar-type* verbs, and (4) the subjunctive in noun clauses. Each video (one video per structure; the format of the videos will be discussed in Sections 4.3 and 4.4) additionally included target vocabulary words. We focused on these four topics 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 topics covered in the course (e.g., *por/para* 'for/to'), only half-day of grammar instruction was included in the syllabus.

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. They were fifth-semester intermediate learners of Spanish, and participated in the study for credit as part of one of their course requirements.[4] The course contained 12 sections, which were quasi-randomly assigned to one of three conditions: a No Captions group (Control); a Captions + TE Vocabulary (Vocabulary) group; and a Captions + TE on the grammatical features group (Grammar) (see Table 2.1 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).

In total, there were six instructors assigned to the twelve sections of the Spanish grammar course. The quasi-random allocation procedure worked as follows: (1) for instructors who were teaching three sections (a total of three instructors), one of each of their sections was assigned to a different condition in order to control for teaching style; (2) the three remaining sections, which were taught by three different instructors, were randomly allocated to Control, Vocabulary or Grammar conditions.

[4] Participants were fifth semester learners of Spanish or had received a high score in their Advanced Placement Spanish course in high school.

Table 2.1. Descriptive Statistics.

Group	N subjects	Age Range		Mean Age (SD)	Gender	
		Min.	Max.		Females	Males
Control	63	17	29	19.02 (1.6)	36	24
Vocabulary	59	17	28	18.69 (1.6)	36	23
Grammar	54	18	23	18.61 (0.9)	38	16

Note. Three participants in the Control group did not specify their gender.

4.2. Written instruments

4.2.1. Language History Questionnaire. Participants completed a Language History Questionnaire (Li, Zhang, Tsai, & Puls, 2013), which included basic demographic questions about their age, gender, and education, and more thorough questions about their experience with different languages.

4.2.2. Spanish vocabulary proficiency test. The Lextale-ESP (Izura, Cuetos & 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 recognized as Spanish words. As recommended by Lemhöfer & Broersma (2012) and Brysbaert (2013), the test was scored using the following formula:

$$\text{Score} = \text{N ‘yes’ to words} - 2 * \text{N ‘yes’ to nonwords.}$$

This scoring formula penalized for guessing behavior, 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 is Izura, Cuetos & Brysbaert (2014) is $\alpha = 0.96$ (N = 90).

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

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).

Table 2.2. 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.

4.2.4. Immediate posttests

Vocabulary recognition test. Participants were tested on their recognition of target vocabulary (see Table 2.2). 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.[5] 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

[5] 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).

was *alberca* “pool”, or *frentos*, when the correct form was *frenos* “braces”. Synonyms not presented in the movie 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 and 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).

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).

Figure 2.1. *Representative slides 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.*

4.3. Grammar Lesson Videos

For each of the four target grammatical structures, a short grammar video lesson was created. Each video lesson summarized 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 2.1). In each practice exercise, participants were presented with a question on the target structure. They were given ten seconds to work through the question on their own and subsequently were provided with the correct answer. During each lab session, the grammar lesson videos were presented prior to the presentation of the animated videos. These grammar lesson videos were designed for the purposes of this study exclusively.

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 and rules, 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. We used a Marantz Pmd620 digital recorder and Shure WH20 head-mounted microphones to conduct the recordings. 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 randomization in each of the videos.

The animated videos were created using Nawmal (www.nawmal.com), an animation program 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 is then automatically lip-synched to fictional characters. The Nawmal software also supports the inclusion of gestures as the characters go through their dialogue, as well as camera movements (e.g., close-ups, panning, dollying), 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, Vocabulary, or Grammar). For each video, captions were added using SRT Edit Pro (http://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 artificial videos, students responded to an exit survey, which consisted of open-ended questions about the usefulness of the animated videos as well as their feedback for the improvement of these videos. A total of 143 students out of the 176 (81%) who participated in the study completed the survey. Overall, we received more positive (total = 126 (83%)) than negative comments (total = 26 (17%)) from students about the usefulness of our 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 2.2 for the breakdown of these target words by session). The target vocabulary chosen for the experiment were either low-frequency

words taken from the NIM Frequency database[6] (Guash, Boada, Ferré & Sánchez - Casas, 2013), or regional vocabulary words to which participants would have only been exposed if they 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 vocabulary words as there were grammar rules being targeted. For instance, for the preterite and imperfect session, there were seven vocabulary targets, the same number of grammar rules presented in the video. Each of the target vocabulary 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 escriba: Curso avanzado de gramática y composición* (Cantelis Dominicis & 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 verb instances, appeared in the animated video.

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 characterize past actions as having a definitive beginning and endpoint (e.g., *caminé* 'I walked'), whereas imperfect forms characterize past actions or states

[6]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.

being viewed as in progress (e.g., *caminaba* 'I was walking / I used to walk'). As noted in Liskin-Gasparro (2000), tense-aspect morphological forms such as the preterite and imperfect differ in their frequency distribution in the input received by L2 learners of Spanish, and thus lead to infrequent exposure of the contrast of these forms. However, as a motivating point for our study, Blyth (2005) asserts that such grammatical forms can benefit from interventions that "render surface forms more frequent and more salient, thereby allowing the learner to focus on form in a meaningful context" (p. 213).

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 (i.e., preterite verbs were accomplishments and achievements, whereas imperfect verbs were activities and states). (see Supplementary Materials; Appendix A; Tables 2.A1-A2, for the complete selection of rules and verbs, and Appendix B1 for an excerpt of the script).

Session 2: Ser and estar. Contrary to the English language which only has one copula verb, 'to be', the Spanish language has two forms, *ser* and *estar*. The standard usage of these forms requires learners to understand the lexical semantic properties that differentiate them. Previous research on the *ser* and *estar* distinction has shown that the target-like usage of these forms is characterized by distinct developmental stages whereby initially, learners omit the usage of both copula verbs, followed by an overgeneralization of the *ser* form, and finally, the proper distinction of the two forms in different contexts at more advanced stages of learning (e.g., *estar* with locatives and *estar* with adjectives denoting condition; Geeslin 2003; VanPatten, 1987). 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 internalize during interlanguage development.

For *ser* and *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; Appendix A; Tables 2. A3-A4, for the complete selection of verbs, and Appendix B2 for an excerpt of the script). 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.

Session 3: Gustar–type verbs. L1 learners' mastery of the *gustar*-type verb construction is considered especially challenging given the marked differences between its English counterpart 'to like' (e.g., Cerezo, Caras & Leow, 2016):

> "Despite their closeness in meaning, these predicates exhibit a divergent syntactic behavior: whereas 'like' codes as subject the entity that experiences a certain feeling, and as object the stimulus responsible for that feeling, gustar expresses the experiencer though an indirect object (or dative) and the stimulus through the subject" (Vázquez Rosa, 2006, p. 1).

For instance, in English, it is standard to construct a sentence that has the subject/experiencer "liking" a direct object (e.g., 'I like red roses'). However, in Spanish a different construction is used: *A mí me gustan las rosas* 'Red roses are pleasing to me', whereby the subject of the sentence is what in English would be considered the direct object, *rosas* 'roses'. The conjugated verb in the Spanish construction depends on whether the subject (i.e., *rosas*) is singular or plural. This L2 syntactic ordering poses challenges for L1 English speakers because it diverts from the canonical subject-verb-object (SVO) word order pattern found in English, instead favoring an

OVS word order as its most frequent syntactic pattern (VanPatten et al., 2009). One common observation in L2 Spanish acquisition research is that, at beginning stages, learners tend to interpret the subject as the first nominal feature in a sentence (e.g., Lee and Malovrh, 2009; Seibert Hanson & Carlson, 2014; VanPatten, 1996). In the case of *gustar*-type verb structures, processing of the initial noun phrase or preverbal object pronoun as the subject can lead to a non-standard conjugation of the main-clause verb. However, there is evidence that target-like processing of OVS structures of this type can be promoted by instructional techniques in which the "connection between form and meaning is made virtually unequivocal and unavoidable" (DeKeyser & Prieto Botana, 2013, p. 456; see also Sanz & Morgan-Short, 2004; VanPatten & Oikkenon, 1996). Whereas most 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 & 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 B3 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 &Prieto Botana, 2013, p.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, 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; see Appendix A, Tables 2.A5-A6 for the complete selection of rules and verbs; see Appendix B4 or an excerpt of the script). Twelve indicative sentences were included as fillers (see Supplementary Materials; Appendix A, Table 2.A7).

A

B

C

Figure 2.2. *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.*

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: The control version of the videos did not include captions.
- Vocabulary: The vocabulary version included captions that provided learners with TE on the target vocabulary via bold and yellow text.
- Grammar: For each of the four target structures, the grammar version included captions that provided learners with TE, again via bold and yellow text on the target grammatical features (See Figure 2.2 for an illustration of these three condition types).

In the present study, captions were inserted within a black background instead of superimposed over the video image. We used boldfacing and color (yellow text) to make the target vocabulary and grammar forms more salient. For the preterite and imperfect, *ser* and *estar*, and *gustar*–type verb structures, we highlighted the relevant verbs, whereas for the subjunctive, we highlighted the phrase containing the verb in the main clause (which can act

as a cue to the subjunctive), the conjunction *que* 'that', and the subordinated subjunctive verb (See Supplementary Material, Appendix C, Figure 2.C1 for an illustration).

Table 2.3. 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.

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] During the first week of class, the learners additionally filled out the web-based Language History

[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 make-up session.

Questionnaire through the LHQ 2.0 interface available on the Penn State Brain Language and Computation Lab website.[8]

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 and imperfect forms, (2) *ser* and *estar*, (3) *gustar-type* verbs, and the (4) subjunctive in noun clauses. Due to the curricular constraints of the grammar course and the common syllabus designed for all 12 sections, it was therefore not possible to counterbalance the presentation of the four grammar topics. During each session, the two experimenters met with the learners and instructors on their assigned class day and time, in a pre-assigned computer classroom. The experimental protocol was computerized and made available to each participant through the Canvas Learning Platform (https://www.canvaslms.com/), which allows for the creation of multimedia surveys. During each experimental session, learners were presented with the grammar lesson video about the target form, followed by the corresponding animated video manipulated for one of three conditions: no captioning was provided (Control group); target vocabulary was highlighted via TE (Vocabulary group); or grammatical features were highlighted via TE (Grammar group). Instructors were asked not to assign readings or homework on the target material prior to the experimental phases.

Following the grammar video lessons and the animated videos, participants completed four different tests, which examined their recognition and production (translation) of the target vocabulary and grammar. Each session lasted approximately 50 minutes. Similar versions of the

[8] The Language History Questionnaire can be accessed online through the Penn State Brain Language and Computation Lab website: http://blclab.org/language-history-questionnaire/

grammar-translation tests were administered by the learners' instructors two weeks after the treatment in order to measure retention of the grammar structures over time (See Table 2.3 for a summary of the procedure).

4.6. Statistical Analysis

Statistical analyses were conducted using R Studio version 1.0.143 (R Studio Team, 2015). The data were analyzed by generalized linear models and multilevel generalized linear regression models utilizing the glm() and glmer() functions within the lme4 package in R (Bates, Maechler, & 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 measures were proportion of trials correct, with GROUP (Control, Vocabulary and Grammar) as the predictor term. The week 1 VOCABULARY PROFICIENCY test was additionally included as a fixed variable 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 modeling the probability of a correct trial), with GROUP (Control, Vocabulary and Grammar), 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 modeling procedure, following the advice of Barr et al. (2013). This modeling 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 versus imperfect session was treated as missing data because they saw the animated video more than once (Control = 15; Vocabulary = 16; 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 2.4. Means, standard deviations (SD), and 95% Confidence Intervals (CI) for the Vocabulary and Grammar Proficiency Tests.

Group	Mean	SD	95% CI
	Group Accuracy scores for Vocabulary Proficiency		
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]
	Group Accuracy scores for Grammar Proficiency		
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]

Table 2.4 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 2.5. Vocabulary Recognition result summary: coefficient estimates β, standard errors SE (β), associated Wald's z-score (= β/SE(β)) 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)	0.774	0.049	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 ***

5.2. Vocabulary Recognition

The Vocabulary recognition post-test data are plotted in the left-hand panel of Figure 2.3 (see also Supplementary Materials; Appendix D; top panel of Table 2.D1 for detailed group accuracy proportion scores). 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 Vocabulary group

participants over the Control and Grammar groups (see the left-hand panel of Figure 2.3). To investigate the effects of captioning, we ran a generalized 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 Vocabulary, $\beta = 1.352$, SE = 0.087, $p < 0.001$ and 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 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 Grammar groups, $\beta = -0.694$, SE = 0.094, $p < 0.001$. Thus, there was an advantage of the Vocabulary group over the Grammar group and the Control Group in their recognition accuracy (see Table 2.5 for complete results summary).

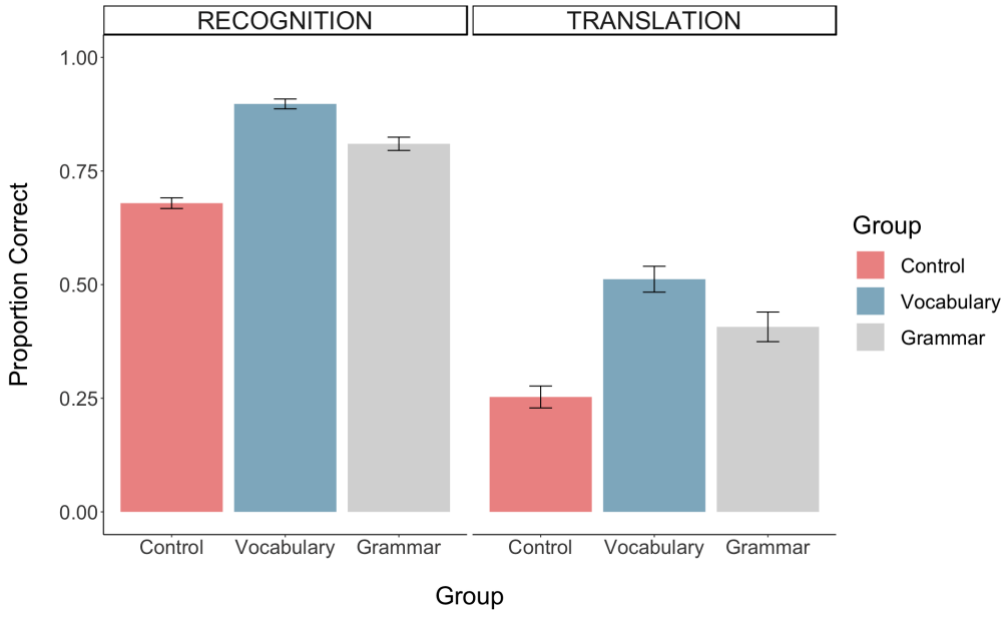

Figure 2.3. *Mean Accuracy Scores for Vocabulary Recognition (left panel), and Vocabulary Translation (right panel). Error bars are 2 standard errors long.*

Table 2.6. Vocabulary Translation result summary: coefficient estimates β, standard errors SE (β), associated Wald's z-score (= β/SE(β)) 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 2.3; and Supplementary Materials; Appendix D; bottom panel of Table 2.D1 for detailed group accuracy proportion scores). 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 Vocabulary, β = 1.034, SE = 0.098, $p < 0.001$, and to Grammar, β = 0.524, SE = 0.105, $p < 0.001$, i.e., both captioned groups were more accurate in their production accuracy. The same model, with Vocabulary as the reference level, revealed a significant negative group effect, compared to Control, β = -1.034, SE = 0.098, $p < 0.001$ and to Grammar, β = -0.510, SE = 0.096, $p < 0.001$, confirming our initial observation of the overall advantage of the Vocabulary group (see Table 2.6 for complete results summary).

5.4. Grammar Recognition

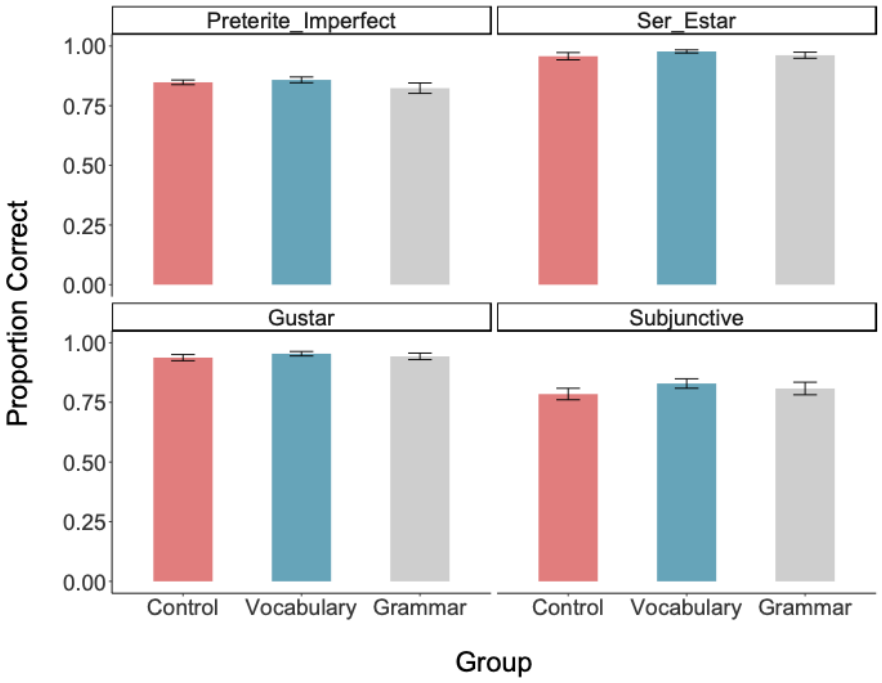

Figure 2.4. *Mean Accuracy Scores for Grammar Recognition by Structure and Group. Error bars are 2 standard errors long.*

Figure 2.4 illustrates the group mean scores as well as the standard errors for all four target grammar structures –the preterite and imperfect, *ser* and *estar*, *gustar*-type verbs and the subjunctive (see also Supplementary Materials; Appendix D; Table 2.D2 for detailed group

accuracy proportion scores). Here, the overall pattern does not suggest any clear group differences within each structure. We ran a generalized 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 Table 2.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 (see also Supplementary Materials; Appendix C; Figure 2C2 for marginal effects plots).

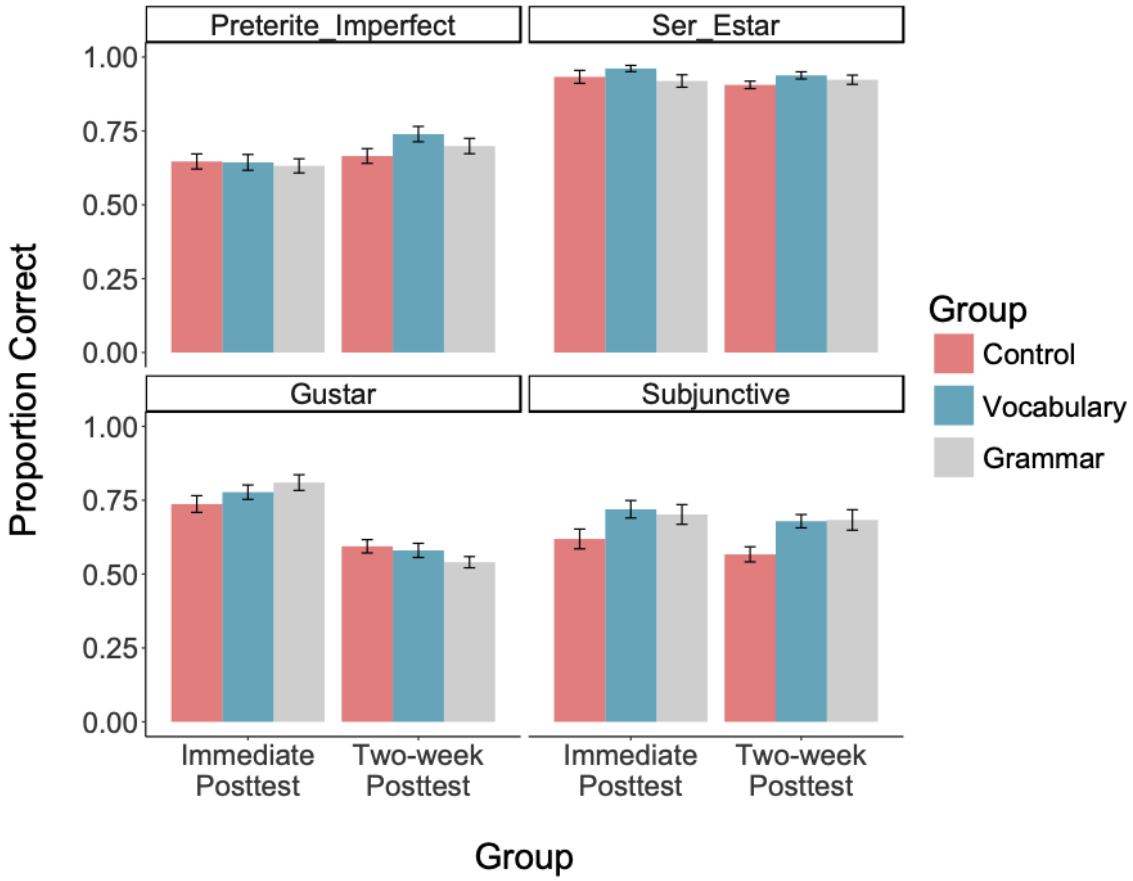

Figure 2.5. *Mean Accuracy Scores for Grammar Translation by Structure, Group and Time. Error bars are 2 standard errors long.*

5.5. Grammar Translation

5.5.1. Immediate Posttest

The left-hand panels of Figure 2.5 plot the group mean scores as well as the standard errors by structure for the Immediate Posttests (see also Supplementary Materials; Appendix 2; top panel Table 2.D4 for detailed group accuracy proportion scores). The data pattern shows varying effects of captioning on production by structure: (1) for the preterite and imperfect, the data does not show any clear differences between groups in their Immediate Posttest scores; (2) for *ser* and *estar*, participants appear to be close to ceiling, with no clear advantage for any group; (3) for *gustar*-type verbs, there appears to be a slight advantage for both the Vocabulary and Grammar

groups over the Control group; finally, (4) for the subjunctive, the data show an advantage for both captioning groups over the Control. To investigate group differences, we ran a generalized 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 Table 2.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 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 vocabulary group also had an effect on the subjunctive, $\beta = 0.525$, SE = 0.199, $p < 0.01$ (see also Supplementary Materials; Appendix C; left-hand panel of Figure 2C3 for marginal effects plots).

5.5.2. Two-week Posttest

The right-hand panels of Figure 2.5 plot the group mean scores as well as the standard errors by structure for the Two-week Posttests (see also Supplementary Materials; Appendix 2; bottom panel Table 2.D4 for detailed group accuracy proportion scores). Again, the data pattern shows varying effects of captioning on production by structure: (1) for the preterite and imperfect, a slight advantage is observed for the Vocabulary group; (2) for *ser* and *estar*, again participants appear to be close to ceiling, with no clear advantage for any group; (3) for *gustar*-type verbs, there appears to be a slight advantage for both the Control and Vocabulary groups; finally, (4) for the subjunctive, again, the data show an advantage for both captioning groups

over the Control group. To investigate group differences, we ran a generalized linear mixed effect model which included fixed effects of grammar proficiency, and our main variables of interest: GROUP, and STRUCTURE as predictor terms, with random intercepts for SUBJECTS. We followed the same 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 only (see Table 2D6 for the full summary). Our follow-up models revealed significant effects when comparing the Control group against the 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 vocabulary group also had an effect on the subjunctive, $\beta = 0.528$, SE = 0.184, $p < 0.01$ (see also Supplementary Materials; Appendix C; right-hand panel of Figure 2C3 for marginal effects plots).

To summarize the grammar results, the immediate posttest data show significant effects of captions + grammar TE on *gustar*-type verbs and on the subjunctive. The results for the two-week posttest reveal that these effects were maintained only for 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 vocabulary words. There was an effect of vocabulary TE on both the recognition and production scores. This is evidenced by the

advantage of the Vocabulary group over both the Control and Grammar conditions – 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, facilitates learners' "taking out" of language (Vanderplank, 2016), independent of the type of enhancement. This may be due, in part, to the adaption process described in Vanderplank (2016), whereby learners select "the language attended to for their own purposes" Vanderplank (2016, p. 239). In this case, the unfamiliarity of the target vocabulary could have led the learners in the 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). Specifically, 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 words 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. The advantage of full captions + TE vocabulary additionally highlights the role of salience and attention as essential factors in L2 learning. Specifically, 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 experimental 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 & 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.

Our results are also in line with the large body of previous work investigating the role of modality of input presentation in L2 acquisition, where it is generally the case that attending to target forms is more difficult in aural than written conditions (Cintrón-Valentín & Ellis, 2016; Morgan-Short et al, 2018; VanPatten, 1990; Wong, 2001; Vidal, 2011). They additionally support the research demonstrating that providing learners with multimodal input, i.e., enriching the aural with written or visual cues, can lead to significant advantages in vocabulary acquisition (e.g., Brown, 2008; Jones and Plass, 2002; Webb & Nation, 2017). Webb and Nation (2017) discuss how the use of elaboration techniques, designed to enrich a learner's knowledge of a word "by encountering more aspects of its form, meaning, and use", such as the inclusion of pictures in addition to written text, can in many instances ,"provide a memorable image of the meaning and context of a word"p.73), and thus facilitate acquisition. In line with this notion, it seems that visual captions can thus serve to make target words more salient, allowing learners to

better home in on and extract relevant information from the input thus support better vocabulary acquisition.

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 vocabulary findings, the findings for grammar were mixed. For recognition, no significant differences were found between the groups for any of the structures, whereas for the production data, captioned videos showed an advantage over non-captioned videos for some structures, namely *gustar*-type verbs and the subjunctive in noun clauses – and this was true both for the immediate posttest and the two-week posttest. Overall, the findings from the production task thus suggest that learner knowledge of some structures is more easily enhanced by captioning than others. In Sections 6.2.1 through 6.2.4, we focus on the specific effects for each structure and provide a discussion on how structure-specific characteristics may have modulated their saliency.

6.2.1. Preterite and imperfect

Regarding the preterite and imperfect forms, learners need to understand how to encode past aspectual distinctions morphologically. As described above, this process involves internalizing a set of rules that describe the contexts in which each form is used. In the current study, we included seven rules – three focusing on the preterite, three focusing on the imperfect, and one focusing on their contrast within one sentential context. Additionally, in order to extract these rules from the input, learners needed to analyze each of the sentential contexts containing the

highlighted preterite and imperfect forms (see Bardovi-Harlig, 1998, regarding the importance of narrative context in the acquisition of tense-aspect morphology), without the provision of additional highlighted cues (as we did with the subjunctive), or visual TE focusing specifically on the morphemes used to mark the aspectual contrast. The number of rules learners needed to analyze here, as well as the lack of additional cues to interpretation, could have influenced the result whereby we did not uncover significant differences among the three learner groups for either recognition or production.

6.2.2. Ser and *estar*

For *ser* and *estar,* learners were at ceiling, and no significant group differences were uncovered between the experimental groups. As mentioned in section 4.4.2, the development of the Spanish copula contrast has been shown to follow distinct developmental stages, where the proper distinction of the two forms in different contexts is more prevalent at more at more advanced stages of learning (e.g., VanPatten, 1987). The learners in the current study were intermediate learners of Spanish, and it could be that they already had ample experience with the copula contrast in their L2 (see section 6.2 for a discussion of this limitation). Although little is known about the degree of prior knowledge that learners require in order to benefit from TE manipulations, in their meta-analysis on TE and grammar learning, 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, correct 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 follow this type of construction. Once acquired, learners need only apply the same rule to each verb instance. Thus, in the context of our study, 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.

6.2.4. Subjunctive in noun clauses

As mentioned previously, 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 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 rules underlying the usage of the subjunctive, as described in Farley & McCollam (2004) (see Collentine (2013) for review).

To summarize, we have suggested here that, within the domain of grammar learning in captioning and TE studies, it is important to take into account structure-specific characteristics that may modulate their saliency. For the four structures examined here, we have brought to light the importance of taking into account the discourse context (i.e., the preterite and imperfect), learners' prior knowledge of the structures in question (i.e., ser and estar), the number of rules being thought of during input processing (i.e., gustar-type verbs and the preterite and imperfect), and how highlighting syntactic dependencies or additional contextual cues may facilitate the learning process (i.e., the subjunctive).

6.3. Limitations

Taken together, the findings from these four structures offer preliminary evidence that captioning can, in some cases, be useful tools for L2 grammar development. However, much work remains in terms of fine-tuning the quantity and types of enhancement that would be required for the successful acquisition of the different grammatical forms. For instance, one limitation of our study is that we did not consider the relative influence of different types of TE on grammar and vocabulary learning, and whether different types of TE have differential effects on L2 learning. For instance, LaBrozzi (2016) investigated the effectiveness of six individual types of TE (e.g., capital letters, font size, underlining, bolding) on the learning of the preterite and imperfect, and found that increased font size on aspectual morphemes led to greater form recognition than in a control group. Comeaux and Macdonald (2017) also used TE to facilitate the acquisition and usage of morphological cues (i.e., case marking and verb agreement) in determining actor assignment in an artificial language and found that learners benefitted from TE

in their production of case-marking, but not of verb agreement (see Tolentino & Tokowicz, 2014, for a study with similar findings).

A second limitation 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 might be beneficial to counterbalance the order of grammar structures to determine whether increased positive effects over time (with *gustar*-type verbs especially) may be a product of participants' growing familiarity with the experimental procedure.

A third limitation of our design is that we did not include a pretest prior to conducting the experimental sessions, as in, for example, LaBrozzi (2016). Without a pretest, it is complicated to tease apart any possible confound regarding the gains acquired through the treatment from pre-existing knowledge. In our study, this issue is relevant when trying to interpret the near-ceiling effects for *ser*/*estar* in all three groups. One question that we cannot answer, for example, is whether the learners already knew these structures going into the study, or if the grammar lesson by itself was sufficiently effective. Future studies would benefit from the inclusion of a pretest design in order to discern the effects of prior knowledge from the experimental treatment.

6.4. Future directions

Moving forward, our results underscore the importance of tailoring TE to each target structure 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 video captions on the structures in question, which is the impetus for current ongoing research (Garcia-Amaya et al., submitted). Future studies should additionally consider the interaction of

different configurations of TE video captions with variables such as learner proficiency (see for instance: Muñoz, 2017) and prior knowledge, each of which has been shown to modulate the effects of TE (Han et al., 2008).

Future research on the effects of TE-captioned videos and grammar learning 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. For instance, Han et al. (2008, p. 601) argue that "the majority of the [TE] studies [have] solely invoked so-called acquisition measures for pre- and post-tests." In their view, developmental designs do not adequately measure learners' immediate noticing of perceptually enhanced input or subsequent learning of the attended form based on successful intake. The inclusion of research tools such as eye-tracking (see for instance Lee & Révész, 2018; Montero Perez, Peters & 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

In conclusion, the potential interactions of grammatical structure, learner proficiency, learner background, amounts and types of classroom instruction to support incidental learning, and types of TE markup are substantial. This multi-scale type of investigation is increasingly demanding of studies and participants. We are beginning to see large-scale interdisciplinary investigations of relevant factors in extensive and varied populations of learners using on-line instruction (Alexopoulou, Michel, Murakami & Meurers, 2017; MacWhinney, 2017; Ziegler et al., 2017). The complexity of these interactions leads us to believe that multiple scales of investigation will be necessary, from experimental classroom research of the type we report here, up through "big-data" investigations including machine-learning from large-learner corpora

(Settles, Brust, Gustafson, Hagiwara & Madnani, 2018), as well as multiple incremental online controlled A/B investigations (https://vwo.com/ab-testing/) on web-learning platforms (Kohavi & Longbotham, 2017).

8. References

Alexopoulou, T., Michel M., Murakami A., & Meurers, D. (2017). Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing natural language processing techniques. *Language Learning, 67*(S1), 180-208.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(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*(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*(4), 471-508.

Bardovi-Harlig, K. (2000). *Tense and aspect in Second Language Acquisition: Form, meaning, and use*. Oxford: Blackwell.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language,* 68(3), 255-278.

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

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

Borrás, I., & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *The Modern Language Journal, 78*(1), 61-75.

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

Brysbaert, M. (2013). Lextale_FR a fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica, 53*(1), 23-37.

Canteli Dominicis, M., & Reynolds, J. J. (2014). *Repase y escriba: Curso avanzado de gramática y composición* (7th ed.). St. John's University: New York.

Cerezo , L. , Caras , A. , & Leow , R. P . (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*(2), 265-291.

Cintrón-Valentín, M. C., & Ellis, N. C. (2016). Salience in Second Language Acquisition: Physical form, learner attention, and instructional focus. *Frontiers in Psychology, 7, 1284.*

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*(1), 3-42.

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. (2018). Determining the effectiveness of visual input enhancement across multiple linguistic cues. *Language Learning, 68*(1), 5-45.

DeKeyser, R. M. (2005). What makes learning second language grammar difficult? A review of issues. *Language Learning, 55*(1), 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.

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. New York, NY: Cambridge University Press.

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*(2), 164-94.

Ellis, N. C. (2017). Salience in usage-based SLA. In S. Gass, P. Spinner, & J. Behney (Eds.), *Salience in Second Language Acquisition* (pp. 21-39). New York: Routledge.

Farley, A. P., & McCollam, K. (2004). Learner readiness and L2 production in Spanish: Processability theory on trial. *Estudios de Lingüística Aplicada, 22*(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*(3), 239-258.

Gass, S. M., Spinner, P., & Behney, J. (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*(4), 703-764.

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), 1-50.

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods, 45*(3), 765–771.

Gudmestad, A. (2012). Acquiring a variable structure: An interlanguage analysis of second-language mood use in Spanish. *Language Learning, 62*(2), 373-402.

Han, Z., Park, E. S., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics, 29*(4), 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*(1), 75-96.

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica, 35*(1), 49-66.

Jones, L., & Plass, J. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal, 86*(4), 546-561.

Jung, U. O. (1990). The challenge of broadcast videotex to applied linguistics. *IRAL-International Review of Applied Linguistics in Language Teaching, 28*(3), 201-220.

Kohavi. R., &, Longbotham, R. (2011). Online controlled experiments and A/B tests. In C. Sammut & G.I. Webb (Eds.), *Encyclopedia of Machine Learning and Data Mining*. Springer, Boston, MA.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149-174.

LaBrozzi, R. M. (2016). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research, 20*(1), 75-91.

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*(3), 307-331.

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, B., & 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. *The Modern Language Journal, 102*(3), 557-577.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods, 44*(2), 325-343.

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*(3), 673-680.

Liskin-Gasparro, J. (2000). The use of tense-aspect morphology in Spanish oral narratives: Exploring the perceptions of advanced learners. *Hispania, 83*(4), 830-844.

MacWhinney, B. (2017). A shared platform for studying second language acquisition. *Language Learning, 67*(S1), 254-275.

Markham, P. L. (1989). The effects of captioned television videotapes on the listening comprehension of beginning, intermediate, and advanced ESL students. *Educational Technology, 29*(10), 38-41.

Markham, P. L. (1993). Captioned television videotapes: Effects of visual support on second language comprehension. *Journal of Educational Technology Systems*, *21*(3), 183-191.

Markham, P. L. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals, 32*(3), 321-328.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41*(3), 720-739.

Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology, 18*(1), 118-141.

Montero Perez, M., Peters, E., & Desmet, P. (2015). Enhancing vocabulary learning through captioned Video: An eye-tracking study. *Modern Language Journal*, *99*(2), 308-328.

Morgan-Short, K., Marsden, E., Heil, J., Issa, B., Leow, R. P., Mikhaylova, A., Mikołajczak, S., Moreno, N., Slabakova, R., & Szudarski, P. (2018). Multi-site replication in SLA research: Attention to form during listening and reading comprehension in L2 Spanish. *Language Learning,68*(2), 392-437.

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

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.

Price, K. (1983). Closed-captioned TV: An untapped resource. *MATESOL Newsletter, 12*(2), 1-8.

RStudio Team (2015). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA. URL: http://www.rstudio.com/.

Sanz, C., and Morgan-Short, K. (2004). Positive evidence versus explicit rule presentation and explicit negative feedback: A computer-assisted study. *Language Learning, 54*(1), 35-78.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Seibert Hanson, A., & Carlson, M. T. (2014). Global effects: The roles of first language and proficiency in L2 processing of Spanish clitics. *Language Learning, 64*(2), 310-342.

Settles, B., Brust, C., Gustafson, E., Hagiwara, M., & Madnani. N. (2018). Second language acquisition modeling. In *Proceedings of the NAACL-HLT workshop on innovative use of NLP for Building Educational Applications (BEA).* ACL.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*(2), 165-179.

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology, 14*(2), 50-73.

Tolentino, L. C., & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning, 64*(2), 279-309.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*(2), 183-203.

Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language teaching, 43*(1), 1-37.

Vanderplank, R. (2016). 'Effects of' and 'effects with' captions. How exactly does watching a TV programme with same-language subtitles make a difference to language learners? *Language Teaching, 49*(2), 235-250.

VanPatten, B. (1987). Classroom learners' acquisition of *ser* and *estar*: Accounting for Developmental Patterns. In B. VanPatten, T. Dvorak, & J. Lee (Eds.), *Foreign Language Learning: A Research Perspective* (pp. 61-75). Rowley, MA: Newbury House.

VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 12*(3), 287-301.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood, NJ: Ablex.

VanPatten, B., & Oikkenon, S. (1996). Explanation vs. structured input in processing instruction. *Studies in Second Language Acquisition, 18*(4), 495-510.

VanPatten, B., Inclezan, D., Salazar, H., & Farley, A. P. (2009). Processing instruction and dictogloss: a study on object pronouns and word order in Spanish. *Foreign Language Annals, 42*(3), 557-575.

Vázquez Rozas, V. (2006). Gustar-type verbs. In J. Clancy Clements, & J. Yoon (Eds.), *Functional approaches to Spanish syntax. Lexical semantics, discourse and transitivity* (pp. 80-114). New York: Palgrave MacMillan.

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

Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition, 30*(1), 79-95.

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford: Oxford University Press.

Winke, P. M., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 14*(1), 65-86.

Winke, P. M., Sydorenko, T., & Gass, S. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *Modern Language Journal, 97*(1), 254-275.

Wong, W. (2001). Modality and attention to meaning and form in the input. *Studies in Second Language Acquisition, 23*(3), 345-368.

Wulff, S., & Ellis, N. C. (2018). Usage-based approaches to SLA. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual Cognition and Language: The state of the science across its subfields* (pp. 37-56). Amsterdam: John Benjamins.

Ziegler, N., Meurers, D., Rebuschat, P., Ruiz, S., Moreno-Vega, J. L., Chinkina, M., Li, W., & Grey, S. (2017). Interdisciplinary research at the intersection of CALL, NLP, and SLA: Methodological implications from an input enhancement project. *Language Learning, 67*(S1), 210-232.

Appendix A.

Selection, placement and randomization of target Rules, verb instances, and vocabulary by Structure

Appendix A1. The Preterite and Imperfect

*Table 2.A1.*Target Rules and examples for the preterite and imperfect.

Structure	**Rule**	**Example**
Preterite	A situation that occurred once	Mi familia hizo (preterite) un viaje a Florida en 1990./My family went on a trip to Florida in 1990.
Preterite	Subsequent actions	Primero tuvo (preterite) un hijo y después estudió (preterite) leyes./First she had a son and then she studied law.
Preterite	Precise/Exact Actions	Conocí a Carla en el 2015./I met Carla in 2015.
Imperfect	Occurred repeatedly in the past	Cuando era niño viajábamos (imperfect) a Disney todos los veranos./When I was Young we would travel to Disney every Summer.
Imperfect	Two simultaneous actions in the past	Ella viajaba (imperfect) mientras yo estudiaba (imperfect)./She would travel while I studied.
Imperfect	Actions that are not as precise	Yo conocía (imperfect) algunos datos sobre la caída del Imperio Romano, pero Mercedes no./ I knew some facts about the fall of the Roman Emprire, but Mercedes did not.
Both	Contrast between preterite and imperfect	Cuando la vi (preterite) ella hacía (imperfect) la comida./When I saw her she was preparing dinner.

Table 2.A2. Rules and Verb Instances for the preterite and imperfect.

	PRETERITE	PRETERITE	PRETERITE	IMPERFECT	IMPERFECT	IMPERFECT	BOTH
Target Rules	Occurred once	Subsequent actions (preterite & preterite):	Precise Actions	Occurred repeatedly in the past	Two simultaneous actions in the past (imperfect & imperfect)	Not as precise	CONTRAST
Target verbs	encontrar	tener/estudiar	entender	navegar	correr/caminar	conocer	ver/buscar
	reconocer	entrar/preguntar	vender	viajar	completar/cuidar	escribir	preparar/llegar
	conseguir	casar/comprar	morir	desear	trabajar/charlar	utilizar	descubrir/dirigir
	ganar	graduar/finalizar	aprender	nadar	bailar/cantar	saber	salir/llover

Table 2.A3. Target Rules and examples for the ser and estar script.

Structure	**Rule**	**Example**
Ser	Trades and professions or social groups	Marta **es** profesora./Marta is a profesor.
Ser	Origin or nationality	**Soy** de Michigan./I am from Michigan
Ser	Permanent or Innate quality	Luis **es** alto./Luis is tall.
Estar	Location	Daniel **está** en Ann Arbor./Daniel is in Ann Arbor.
Estar	Estar + past participle; refers to a state or condition resulting from a previous action	**Estamos** preocupados porque hemos sacado malas notas./We are worried because of our bad grades.
Estar	Estar + adj.; Condition or state of the subject	Anita **está** triste./Anita is sad.

Table 2A4. Rules and verb instances for ser and estar.

Target Rules	SER - professions	SER - origin or nationality	SER – permanent or innate quality	ESTAR - location	ESTAR + past participle	ESTAR + adj. Condition or state of the subject
Target verbs	Soy	Es	Soy	Estoy	Estoy	Estoy
	Soy	Es	Soy	Estás	Estoy	Estoy
	Eres	Es	Eres	Está	Estoy	Está
	Soy	Es	Es	Está	Estás	Estás

Table 2.A5. Target Rules and examples for the subjunctive in noun clauses.

Rule	**Example**
Wishes, desires, Imperatives	**Quiero** que llames a tu hermana mañana./I want you to call your sister tomorrow
Emotion	**Me alegro** de que vengas a visitarme en mi cumpleaaños./I am happy that you come visit me on my Birthday.
Doubt, Denial, Disbelief	**Dudo** que ganes un premio en la Feria Científica./I doubt that you will win a prize in the Science Fair.
Impersonal Observations	**Es possible** que vaya a la fiesta./It is posible that I will go to the party.
Recommendations	**Recomiendo** que leas el capítulo antes del examen./I recommend that you read the chapter before the test.

Table 2.A6. Rules and Verb Instances for the subjunctive in noun clauses.

Target Rules	Wishes, desires, Imperatives	Emotion	Doubt	Impersonal Observations	Recommendations
Target structures	**Quiero** que vean	**Me alegro** de que quieras	**Dudo** que encuentren	**Es importante** que hable	**Recomiendo** que tenga
	Deseo que haga	**Me enfada** que llame	**No creo** que vayamos	**Es posible** que envie	**Sugiero** que piense
	Quiero que vengas	**Me alegro** de que trabajes	**Dudo** que haya	**Es importante** que entiendas	**Recomiendo** que esperes
	Deseo que enseñes	**Me enfada** que juegue	**No creo** que sea	**Es posible** que entre	**Sugiero** que dejes

Note. The trigger verbs in the main clause are bolded, and the subjunctive verbs are underlined.

Table 2.A7. Indicative Fillers for the subjunctive in noun clauses.

Location in Script	Part 1	Part 2	Part 3	Part 4
Structures	**creo** que me espera	**me parece** que descartado	**creo** que las construyen	**me parece** que podemos
	veo que va	**afirmo** que queremos	**afirmo** que hemos considerado	**veo** que llegas
	sé que representa	**estoy** seguro de que había recibido	**sé** que quieren	**estoy** seguro se puede
	es cierto que tiene	**es** obvio que yo ando	**es** cierto que se indica	**es** obvio que me darán
	está claro que quieren	**es** verdad que se me olvidan	**es** verdad que consideran	**es**tá claro que te tienen
	es evidente que ella es	**es irrefutable** que hemos trabajado	**es** evidente que caracteriza	**es irrefutable** que nos va a dar

Note.The trigger verbs in the main clause are bolded, and the indicative verbs are underlined.

Appendix B.

Animated Video Scripts by Structure

B1. Excerpt of Animated Video Script for the Preterite and Imperfect Session

PART 1. Introducción:

1. **Ana:** ¿¡Carlos, no me lo puedo creer!? ¿Qué haces aquí?
2. **Carlos:** Ana, hace mucho que no te veo.
3. **Ana:** Has cambiado mucho. Con tu nuevo peinado no te **reconocí [preterite]** .
4. **Carlos:** Y tú también. Ya no tienes los frenos **[vocabulary]** en los dientes.
5. **Ana:** Ay, ¿¡te acuerdas de los frenos**[vocabulary]**?! Tener frenos **[vocabulary]**en los dientes, lo peor. Pero no me puedo quejar. Gracias a los frenos**[vocabulary]** tengo los dientes derechos (Risas)
6. **Carlos:** Hace cinco años que no nos vemos...desde la universidad.
7. **Ana:** Qué divertidos los años de universidad. **Sabía [imperfect]** que te mudaste a la Argentina.
8. **Carlos:** Sí, pero llevo una semana en los Estados Unidos…porque, ¿sabes? El mes pasado **vendí [preterite]** mi apartamento en Buenos Aires.
9. **Ana:** ¿Y por qué has vuelto?
10. **Carlos:** Bueno, **Conseguí [preterite]** un trabajo nuevo en Ann Arbor. Como sabes en el pasado **deseaba [imperfect]** irme lejos para explorar el mundo, pero ahora quiero estar en casa con los míos.

B2. Excerpt of Animated Video Script for the *Ser and Estar* Session

PART 1. *Introducción*

1. **Secretaria:** Buenos días.
2. **Andrés [Llega corriendo porque va tarde]**: Buenos días, me llamo Andrés González. La Sra. Lourdes Molina me citó para una entrevista a las 11:00. ¿**Estoy [estar]** en el piso correcto?
3. **Secretaria: Estás [estar]** en el piso correcto. Yo **soy [ser]** la secretaria de la Sra. Molina. ¿Pero a las 11:30? ¡**Estoy [estar]** sorprendida! Usted ha llegado muy tarde. La Sra. Molina lleva 30 minutos esperándole.
4. **Andrés:** Sí, lo sé. Perdón. Lo siento muchísimo. Con esta lluvia y con un tráfico terrible… imposible encontrar taxi.
5. **Secretaria:** Le preguntaré a la Sra. Molina si tiene tiempo. Todos los candidatos para la posición de publicidad han llegado temprano a la entrevista. ¿**Es [ser] impuntual[vocabulary]**?
6. **Andrés:** No, todo lo contrario, yo no **soy [ser] impuntual [vocabulary].** Siempre llego a tiempo para todo.

7. **Secretaria:** Bueno, pero no se puede ser **impuntual** **[vocabulary]**en una entrevista de trabajo. Voy a hablar con la Sra. Molina pero generalmente no suele atender a gente **impuntual[vocabulary].**
8. **Andrés:** Gracias.

[Transición]

9. **Secretaria:** Andrés, tengo buenas noticias. La Sra. Molina sí podrá entrevistarte.
10. **Andrés:** Ay, estupendo. Muchísimas gracias por su ayuda.

B3. Excerpt from Animated Video Script for the *Gustar*-type verbs Session

PART 3. Cuestiones personales

1. **Lola**: Una cosa, y tú ¿siempre has usado **espejuelos[vocabulary]**?
2. **Pablo:** Sí, desde pequeño he tenido que usar **espejuelos[vocabulary].** Los problemas de visión me vienen de mi madre y a ella de su padre, así que en mi familia todo el mundo lleva **espejuelos[vocabulary]**.
3. **Lola:** Bueno, pues yo igual. Uso **espejuelos[vocabulary]** para mirar la computadora, pero hoy llevo lentes de contacto.
4. **Pablo:** ¡Otra cosa en común, Lola! (Risas)
5. **Pablo:** Lola, tengo una pregunta para ti. ¿te **interesan** los deportes?
6. **Lola:** Bueno, un poco. Pero mira, me voy a sincerar contigo. En realidad me **molestan** los deportes y no sigo ninguno. A veces, asisto a los partidos de fútbol regional de mi hermano..
7. **Pablo:** Mmm, pues me **gustan** mucho los deportes, y lo que más el fútbol. Quién sabe si algún día te veo en algún partido. Suelo asistir a los partidos regionales también.
8. **Lola:** ¿Quién sabe?… *[sonriente]*. Bueno, Pablo, lo he pasado muy bien contigo. Esta cita a ciegas ha ido muy bien.
9. **Pablo:** Pienso igual. No me arrepiento de haber venido. ¿Quisieras repetirla en el futuro?
10. **Lola:** ¡Pablo, qué directo tú! *[risa*

B4. Animated Video Script for the Subjunctive in Noun Clauses Session

PART 3. ***Presentación del Diseño***

1. **Martín:** Aquí tenemos la imagen del salón de recreo con mesa de billar. **Es verdad que los clientes la consideran [indicative]** de gran importancia. Habíamos pensado en una habitación multiusos para todos los miembros de la familia.
2. ***Carmen:*** Interesante. No olvides que los abuelos también se mudarán con ellos y tienen problemas de movilidad.
3. **Martín:** Sí, sí, **le afirmo que hemos considerado [indicative]** incluir un ascensor.
4. **Carmen: Dudo que haya [subjunctive]** espacio para un ascensor en este diseño pero lo veremos ¿Qué hay de la biblioteca? ¿Recuerdas el énfasis en mi lista de detalles? Los clientes quieren una biblioteca con referencias orgánicas y con grandes **estanterías[vocabulary]**.
5. **Martín:** Lo recuerdo perfectamente. Todos en la familia sienten una gran pasión por la lectura y la naturaleza. Así, hemos pensado en construir una biblioteca flotante abierta al jardín. Tendrá multitud de **estanterías[vocabulary]** para todos sus libros. Mira, aquí tengo un antiguo diseño de otra casa ya finalizada: [Mostrar foto]
6. **Carmen:** Mmm…qué buena idea. Y qué atractivas las **estanterías[vocabulary]** por toda la habitación. Mis clientes van a sentirse muy atraídos por este concepto. Martín, **me alegro de que trabajes [subjunctive]** con los deseos de nuestros clientes en mente.
7. **Martín:** Muchas gracias. Estas **estanterías[vocabulary] creo que las construyen [indicative]** en Macao de madera de ébano africano.
8. **Carmen:** Enséñame los dormitorios.
9. ***Martí*n: Es cierto que se indica [indicative]** en la lista la importancia de la oscuridad para el descanso de sus clientes.
10. **Carmen**: Sí, ellos quieren mucha oscuridad en su habitación.

Appendix C.

Additional Figures

Figure 2.C1. *Illustration of the Grammar captioning manipulation taken from the subjunctive session.*

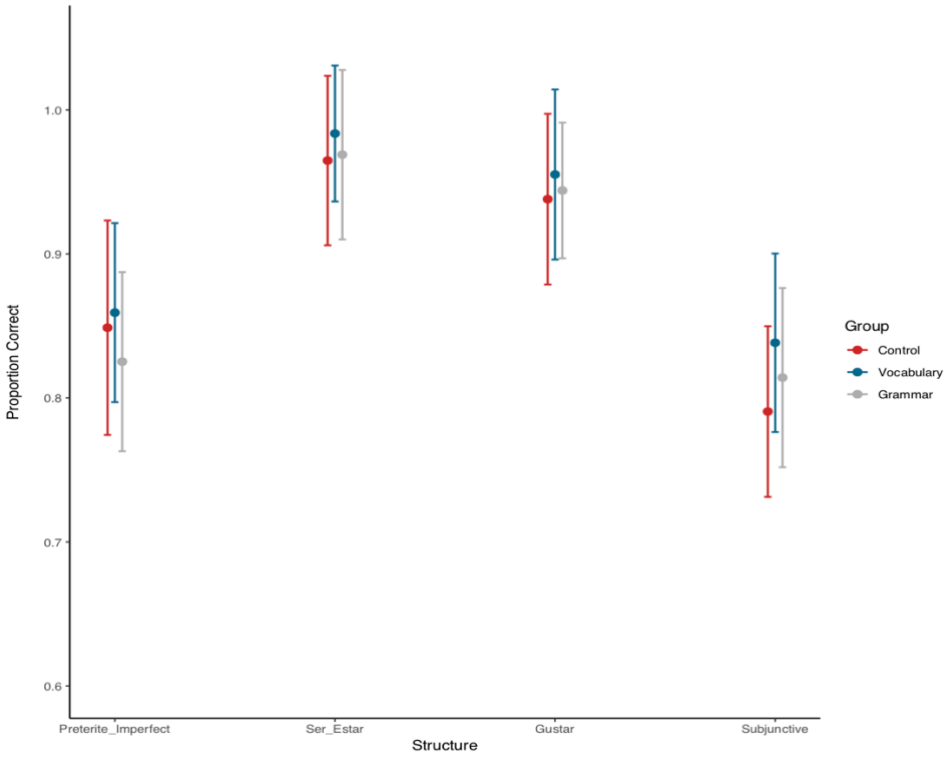

Figure 2.C2. Marginal effects plot for Grammar Recognition. Error bars are Confidence Intervals.

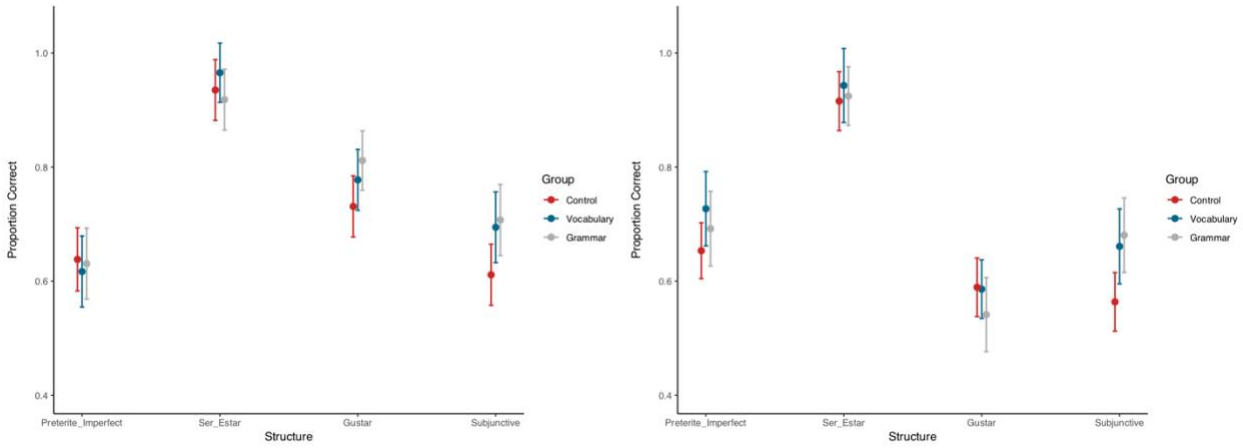

Figure 2.C3. Marginal effects plot for Grammar Translation Immediate Posttest (left-hand panel) and Two-week Posttest (right-hand panel). Error bars are Confidence Intervals

Appendix D.

Additional data tables

Table 2.D1. Group Accuracy Proportion Scores for Vocabulary Recognition and Translation.

Group	Mean	SD	95% CI
	Group Accuracy scores for Recognition		
Control	0.682	0.081	[0.656, 0.702]
Vocabulary	0.898	0.075	[0.876, 0.918]
Grammar	0.810	0.899	[0.781, 0.838]
	Group Accuracy scores for Translation		
Control	0.253	0.169	[0.206, 0.300]
Vocabulary	0.512	0.197	[0.456, 0.568]
Grammar	0.407	0.206	[0.343, 0.471]

Note. CI = confidence interval

Table 2.D2. Group Accuracy Proportion Scores for Grammar Recognition by Structure.

Group	Mean	SD	95% CI
	Group Accuracy scores for the Preterite and Imperfect		
Control	0.848	0.062	[0.835, 0.860]
Vocabulary	0.858	0.069	[0.844, 0.872]
Grammar	0.824	0.109	[0.799, 0.847]
	Group Accuracy scores for Ser y Estar		
Control	0.957	0.106	[0.936, 0.979]
Vocabulary	0.977	0.045	[0.969, 0.986]
Grammar	0.962	0.079	[0.944, 0.979]
	Group Accuracy scores for Gustar		
Control	0.938	0.089	[0.919, 0.955]
Vocabulary	0.954	0.065	[0.941, 0.961]
Grammar	0.943	0.087	[0.924, 0.962]
	Group Accuracy scores for the Subjunctive		
Control	0.785	0.163	[0.752, 0.817]
Vocabulary	0.829	0.137	[0.802, 0.857]
Grammar	0.808	0.161	[0.773, 0.843]

Note. CI = confidence interval.

Table 2.D3. Grammar Recognition result summary: coefficient estimates β, standard errors SE (β), associated Wald's z-score (= β/SE(β)) and significance level p for all predictors in the analysis.

Predictor	Coef. *B*	*SE (β)*	**z**	***p***
(Intercept)	1.752	0.101	17.273	<0.001***
vocabulary	0.097	0.152	0.638	0.524
grammar	-0.146	0.159	-0.920	0.357
ser and estar	2.462	0.472	5.217	<0.001***
gustar	1.209	0.255	4.740	<0.001***
subjunctive	-0.434	0.146	-2.978	<0.01**
grammar proficiency	0.205	0.050	4.099	<0.001***
vocabulary:ser and estar	0.388	0.574	0.676	0.499
vocabulary:gustar	0.223	0.349	0.655	0.512
vocabulary:subjunctive	0.211	0.218	0.969	0.333
grammar:ser and estar	0.231	0.570	0.405	0.686
grammar:gustar	0.303	0.352	0.859	0.390
grammar:subjunctive	0.350	0.223	1.543	0.123

Note. The preterite/imperfect and the control group were the reference levels

Table 2.D4. Group Accuracy Proportion Scores for Grammar Translation by Time and Structure.

Group	Mean	SD	95% CI
IMMEDIATE POSTTEST			
Group Accuracy scores for the Preterite and Imperfect			
Control	0.647	0.155	[0.602, 0.691]
Vocabulary	0.643	0.145	[0.600, 0.687]
Grammar	0.632	0.115	[0.595, 0.668]
Group Accuracy scores for Ser y Estar			
Control	0.933	0.149	[0.890, 0.976]
Vocabulary	0.961	0.069	[0.941, 0.982]
Grammar	0.912	0.128	[0.878, 0.959]
Group Accuracy scores for Gustar			
Control	0.737	0.189	[0.682, 0.792]
Vocabulary	0.777	0.158	[0.729, 0.824]
Grammar	0.809	0.163	[0.758, 0.862]
Group Accuracy scores for the Subjunctive			
Control	0.619	0.219	[0.555, 0.682]
Vocabulary	0.719	0.192	[0.662, 0.777]
Grammar	0.702	0.198	[0.639, 0.765]

TWO-WEEK POSTTEST			
Group Accuracy scores for the Preterite and Imperfect			
Control	0.666	0.148	[0.623, 0.708]
Vocabulary	0.739	0.145	[0.696, 0.783]
Grammar	0.699	0.127	[0.659, 0.739]
Group Accuracy scores for Ser y Estar			
Control	0.906	0.085	[0.881, 0.930]
Vocabulary	0.938	0.079	[0.914, 0.961]
Grammar	0.923	0.096	[0.893, 0.954]
Group Accuracy scores for Gustar			
Control	0.594	0.153	[0.549, 0.638]
Vocabulary	0.580	0.157	[0.533, 0.627]
Grammar	0.540	0.117	[0.503, 0.578]
Group Accuracy scores for the Subjunctive			
Control	0.567	0.173	[0.517, 0.617]
Vocabulary	0.679	0.147	[0.635, 0.723]
Grammar	0.683	0.213	[0.616, 0.751]

Note. CI = confidence interval

Table 2.D5. Immediate Grammar Translation result summary: coefficient estimates β, standard errors SE (β), associated Wald's z-score (= β/SE(β)) and significance level p for all predictors in the analysis.

Predictor	Coef. *B*	*SE (β)*	**z**	***p***
(Intercept)	0.633	0.112	5.660	<0.001***
vocabulary	-0.096	0.169	-0.568	0.570
grammar	-0.036	0.177	-0.205	0.838
ser and estar	3.155	0.435	7.259	<0.001***
gustar	0.462	0.127	3.624	<0.001***
subjunctive	-0.166	0.149	-1.110	0.267
grammar proficiency	0.235	0.058	4.034	<0.001***
vocabulary:ser and estar	0.501	0.554	0.904	0.366
vocabulary:gustar	0.342	0.196	1.745	0.081
vocabulary:subjunctive	0.621	0.225	2.759	<0.01**
grammar:ser and estar	-0.284	0.541	-0.525	0.599
grammar:gustar	0.533	0.206	2.582	<0.01**
grammar:subjunctive	0.539	0.235	2.298	<0.05*

Note. The preterite/imperfect and the control group were the reference levels

Table 2.D6. Two-week Grammar Translation result summary: coefficient estimates β, standard errors SE (β), associated Wald's z-score (= β/SE(β)) and significance level p for all predictors in the analysis.

Predictor	Coef. *B*	*SE (β)*	**z**	***p***
(Intercept)	0.697	0.099	6.985	<0.001***
vocabulary	0.397	0.152	2.605	<0.01**
grammar	0.209	0.158	1.324	0.185
ser and estar	1.695	0.207	8.173	<0.001***
gustar	-0.311	0.126	-2.465	<0.05*
subjunctive	-0.412	0.126	-3.267	<0.01**
grammar proficiency	0.136	0.039	3.477	<0.001***
vocabulary:ser and estar	0.023	0.304	0.077	0.939
vocabulary:gustar	-0.474	0.188	-2.517	<0.05*
vocabulary:subjunctive	0.097	0.192	0.507	0.612
grammar:ser and estar	0.062	0.305	0.204	0.838
grammar:gustar	-0.409	0.194	-2.108	<0.05*
grammar:subjunctive	0.328	0.199	1.646	0.099

Note. The preterite/imperfect and the control group were the reference levels

CHAPTER 3. Assessing Textually Enhanced Caption Designs for their Effects on L2 Grammar Uptake

1. Introduction

1.1. Overview

With the proliferation of multimedia language-learning materials such as online videos, television programs, and mobile language-learning platforms (e.g., Duolingo), it is of growing interest to investigate how traditional instructional techniques such as Input Enhancement (IE) can be effectively integrated with such multimedia approaches to language learning. The current study investigates one multimedia resource, namely that of captioned videos that include visually enhanced intralingual subtitles where the text is presented in the same language as the audio. Specifically, we examine the effect of Textual Enhancement (TE) – a type of IE (Sharwood Smith, 1993) – on the L2 acquisition of three elements of Spanish morphosyntax (i.e., *gustar-type* verbs, preterite/imperfect contrast, and subjunctive in noun clauses). Our experimental design focuses on learner uptake of these three structures, understanding uptake as a learner's immediate recall of the target stimulus (with or without relevant textual enhancement).

In order to investigate the effectiveness of TE in aiding learners' reproduction of target grammatical forms, we implemented an adapted version of the Elicited Imitation (EI) task (e.g., Tracy-Ventura, Ortega, & Norris, 2014; Gaillard & Tremblay, 2016) in a written recall format, which we refer to as a 'Written Elicitation Imitation' task (WEI). The broad aim of this study is to assess the best way to highlight relevant aspects of a grammatical structure; our results are

intended to inform L2 researchers and instructors of the most practical and efficient means of building pedagogically useful explanations of complex grammar in the L2.

1.2. Rationale

The acquisition of grammar and morphology is one of the most challenging aspects of L2 acquisition (DeKeyser, 2005). The Second Language Acquisition (SLA) literature provides various accounts of this difficulty, including age effects and critical periods for language acquisition (e.g., Johnson & Newport, 1989; Hartshorne et al., 2018), individual differences (e.g., Dörnyei, 2005), input processing differences between native (L1) speakers and L2 learners (VanPatten, 1996; 2003), as well as the linguistic features of the target grammar constructions themselves, such as frequency (e.g., Ellis, 2002), complexity, and perceptual salience (e.g., Ellis, 2017; Gass, Spinner & Behney, 2017; Goldschneider & DeKeyser, 2001; Larsen-Freeman, 1976). Perceptual salience refers to the intrinsic qualities of a linguistic cue or structure that aid in its prominence and subsequent cognitive processing and learning. There is evidence that the low perceptual salience of certain grammatical features, such as inflectional suffixes that vary based on the tense/mood/aspect of the grammatical context, contributes to L2 learners' difficulty in acquiring them (Cintrón-Valentín & Ellis, 2016; Ellis, 2006; Gass et al., 2017; Goldschneider & DeKeyser, 2001). To mitigate the acquisitional challenge presented by low perceptual salience, SLA researchers have employed IE techniques such as TE, which involves the use of visual manipulations such as color-coding, boldfacing, and underlining to increase the prominence of such cues in the input (Gass et al., 2017; Sharwood Smith, 1993).

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. Multimodal techniques, which aim to enrich the written input using aural and/or visual cues, were missing from the literature until recent work by Lee and Révész (2018) and Cintrón-Valentín, García-Amaya and Ellis (under review). For instance, integrating theoretical principles from the salience and the grammar-learning literatures, Cintrón-Valentín et al. investigated how captioned media could serve as a useful tool for advancing L2 grammar learning. Their study focused on four Spanish grammar structures, showing significant effects of TE-captions on some, but not all target forms. Such results are not unexpected — the grammar-learning literature shows mixed findings on the effectiveness of TE (e.g., Han, Park, & Combs, 2008; Lee & Huang, 2008; Leow & Martin, 2017). To this point, Comeaux and McDonald (2017) suggest that the efficacy of TE may depend on the type of grammar structure under examination, as well as how the TE manipulation is realized.

Weighing these considerations, as well as the inconsistent findings of previous research regarding TE, we designed an innovative experimental methodology to examine the differential effects of TE. Specifically, we compared the effect of tailoring TE on a full lexical entry (e.g., the complete verb form containing target morpheme) to tailoring TE on a target morpheme only (e.g., past-tense suffix, but not the root) along with any relevant grammatical dependencies. In doing so, we offer an analysis of learners' uptake of three grammatical structures in L2 Spanish: *gustar*-type verbs, the preterite/imperfect contrast, and the subjunctive in noun clauses. In order to motivate our experimental design, Section 2 provides an overview of salience and L2 uptake in grammar learning.

2. Background

2.1. Salience, grammar learning, and L2 uptake

The role of salience 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 & 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, p. 21). This is especially relevant for the acquisition of grammar and morphology given the low perceptual salience that characterizes certain inflectional morphemes (e.g., Brown, 1973; Goldschneider & DeKeyser, 2001; for Spanish, see Collentine (2013, p. 454)). In fact, one common observation in the SLA literature is that despite the vast availability of grammatical forms in the input, L2 learners often ignore certain aspects of morphological structure and focus their attentional resources to the meanings of open-class words, such as nouns, verbs, adjectives, and adverbs, during input processing (e.g., Bardovi-Harlig, 1992; Clahsen & Felser, 2006; Schmidt, 2001; VanPatten, 1996, 2003). L2 morphology is more challenging for learners where variant inflectional morphemes are less salient in the input. Additionally, the functional interpretations are less evident than the one-to-one mappings typical for vocabulary (DeKeyser, 2005; Ellis, 2017; Goldschneider & DeKeyser, 2001).

One way of counteracting the effects of low-salience grammatical forms is by providing learners with IE techniques designed to render target structures more salient (Doughty & Williams, 1998; Han et al., 2008; Lee & Huang, 2008; Leow & Martin, 2017; Sharwood Smith, 1993). TE, for example, involves visual manipulations in written input and thus facilitates learners' processing of target grammatical forms (Sharwood Smith, 1993). However, research in this area has yielded inconsistent findings regarding its effectiveness (for a review and meta-analysis see: Han et al., 2008; Lee & Huang, 2008). Some studies suggest that TE is successful in drawing learners' attention to the target forms (Alanen, 1995; Cho, 2010; Izumi, 2002; Jourdenais, 1995; Lee, 2007; Winke, 2013), as well as in learners' subsequent learning of such

forms (e.g., Jourdenais, 1995; Lee, 2007; Shook, 1994), whereas others have found no effect of TE on learning (Izumi, 2002; Leow, 1997, 2001; Leow, Egi, Nuevo, & Tsai, 2003; Overstreet, 1998; Wong, 2003).

More recent research on the effects of TE on L2 grammar acquisition suggest that its efficacy may be modulated not only by the linguistic form in question (e.g., Cintrón-Valentín et al., under review; Comeaux & McDonald, 2017; Leow et al., 2013) but also by the type of TE that is used (LaBrozzi, 2016). 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 (see also Jourdenais et al., 1995; Leeman et al., 1995; Overstreet et al., 1998). Critically, studies of this nature compare the effects of different TE manipulations (e.g., upper-case versus lower case), but do not focus on tailoring the TE to a target morpheme (e.g., past-tense suffix) in comparison to tailoring the TE to a full lexical entry (e.g., the full verb form that contains the target morpheme). As pointed out in Lee and Huang (2008, p. 327), the next logical step in exploring any substantive effect of TE is through the design of studies "that probe the underlying questions of what exactly influences learners' perception of enhanced forms and how the processing of these enhanced forms might facilitate L2 grammar learning." In our view, this step involves a more focused analysis of the effects of enhancement on a target morpheme, emphasizing the appropriate inflectional and functional considerations. To our knowledge, little or no experimental work has been carried out in this specific area of TE research.

Another critical factor to consider in L2 grammar acquisition is the learning of syntactic constituents and their dependencies. A fundamental aspect of language learning and processing is "the ability to track syntactic relationships between words and phrases in a sentence" (Wilson et al., 2018, p.1). These syntactic dependencies range from those that are simpler in their relation (i.e., between adjacent words) to those that are more complex (i.e., between non-adjacent words), and are more cognitively taxing on the L2 learner due to the distance required to process the structural relationship. The learning of non-adjacent linguistic forms has been found to be more difficult than that of adjacent linguistic forms (for a review, see Wilson et al., 2018). Nonetheless, research shows that it can be facilitated by the provision of additional cues aimed at enhancing their perceptual similarity in the face of other sentential elements (e.g., Newport & Aslin, 2004; Pacton & Perruchet, 2008).

The current study seeks to explore and exploit the principles governing L2 grammar learning by assessing which designs of TE are most optimal for focusing learner attention on different linguistic constructions. Specifically, we investigate the possibility of a differential effect of contrasting input-enhanced captions on L2 uptake. Researchers in SLA typically conceptualize 'uptake' based on a learner's output that immediately follows an instructor's feedback/intervention – this output constitutes a reaction in some way to the instructor's original intention, presumably to draw attention to a target linguistic element in the L2 (Slimani, 1992; see also Allwright, 1984). For example, Ellis et al. (2001) operationalize uptake as learner utterances that occur after either instructor feedback or any interlocutor utterance that provides information about a target linguistic feature. In our research, we conceptualize uptake as what learners produce in response to experimental variations of textually enhanced captions that accompanied animated videos. We opted for an uptake study given its methodological

efficiency: the data-collection process occurs in a single session, while still offering insights into learners' immediate noticing and processing of the captioned material.

Furthermore, as argued in Han et al. (2008, p. 601), "the majority of the [TE] studies [have] solely invoked so-called acquisition measures for pre- and post-tests." In their view, developmental designs do not adequately measure learners' immediate noticing of perceptually enhanced input or subsequent learning of the attended form based on successful intake (see also Jourdenais et al., 1995; Leow et al., 2003). Through our experimental design, we address these concerns by measuring learners' accuracy immediately after receiving enhanced input. Our experimental paradigm, therefore, can inform the larger body of research on L2 learning regarding the usefulness of captions for facilitating learner processing of grammatical forms.

In Section 2.2, we offer a more thorough overview of the captioning research, and how captioning can help draw learners' attention to unknown grammatical forms, thereby promoting L2 noticing and learning.

2.2. Captioning research: Overview

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, realizing the potential of this resource for other target populations, educational researchers began investigating the benefits of captioning with the purpose of developing L2 language skills in the hearing population. The early research on captioning primarily focused on determining if captioned video was better than non-captioned video in (i) improving learner comprehension of video content (e.g., Garza, 1991; Markham, 1993, 1999; Price, 1983), and (ii) promoting vocabulary learning (e.g., Huang & Eskey, 1999; Neuman & Koskinen, 1992). Although

vocabulary learning and comprehension have remained the primary focus of the L2 captioning literature, more recent work has focused on uncovering what specific factors might mediate its effectiveness. For instance, there are studies focusing on the differential effects of captioned video across languages, as well as ordering effects resulting from multiple presentations of the captioned videos (Winke, Gass, & Sydorenko, 2010; Winke, Sydorenko, & Gass, 2013), modality effects (e.g., visual and/or aural) in vocabulary acquisition through captioned video (Sydorenko, 2010), the role of age and proficiency on learners' reading behavior (Muñoz, 2017), and the effect of textual enhancement in the captioning line (Montero Perez, Peters, Clarebout, & Desmet, 2014). Montero Perez et al. (2014), for example, reported a significant positive effect of TE-captioned video, in combination with multimodal input, on L2 vocabulary learning.

On the benefits of captioning for L2 vocabulary learning, some researchers have suggested that the presentation of multimodal input (i.e., 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 & Lafayette, 1994, p.70; see also Garza, 1991; Webb & Nation, 2017). To this point, Vanderplank (2016) and Winke et al. (2010) suggest that captioning can help draw learners' attentional focus to unknown word forms and promote subsequent noticing and learning through repeated exposure. This notion is consistent with foundational theories in SLA which stress that attention is central to successful L2 acquisition (e.g., Gass et al., 2017; Schmidt, 2001; Tomlin & Villa, 1994). Schmidt's (2001, p. 30) Noticing Hypothesis, for instance, holds that conscious attention to linguistic forms in the input (referring to grammatical input especially), is an essential precondition to learning.

In our study, we investigated the effects of TE within the captioning line on L2 grammar uptake through three experimental conditions: a No-Captions Control (NC) condition which presented L2 audio but no material in the captioning line; a TE1-condition which presented target verbs highlighted in their entirety; and a TE2-condition which presented highlighting of the critical morphological and grammatical cues, and their relations, in the captioning line. We targeted three grammatical topics (see Section 4.5) and created one unique video per target construction.

2.3. Research Questions

The current study aimed to extend previous research on captioning and L2-grammar learning by focusing on three research questions:

(1) What is the relative effect of morpheme-enhanced TE, lexical item-enhanced TE, or no TE on the uptake of *gustar*-type verbs in L2 Spanish?

(2) What is the relative effect of morpheme-enhanced TE, lexical item-enhanced TE, or no TE on the uptake of the preterite/imperfect in L2 Spanish?

(3) What is the relative effect of morpheme and dependency-enhanced TE, lexical-item and dependency-enhanced TE, or no TE on the uptake of the subjunctive in noun clauses in L2 Spanish?

3. Method

3.1. Measuring Uptake

We used an adapted version of the Elicited Imitation (EI) task to assess the effect of TE on each

of the three grammar structures. The EI task is well-known in SLA as a measure of language proficiency (Tracy-Ventura et al., 2014). Typically, learners are instructed to listen to a sentence and repeat it verbatim, to the greatest extent possible (the repetition is generally oral, but there are also studies that have considered written imitation, see Vinther, 2002). One assumption underlying the EI method is that learners should be more successful in repeating sentences that contain grammatical structures that are familiar or known to them, and less successful in those where the structure far exceeds their knowledge. In our study, we used a Written Elicited Imitation (WEI) task to investigate whether TE draws learners' attention to relevant parts of the input so to allow them to reproduce the target grammatical structures in a written recall format.

3.2. Participants

A total of 31 English-speaking L2 learners of Spanish were recruited from two summer-term Spanish courses at a large Midwestern University. These courses were designed for students who had just completed the university's obligatory intermediate Spanish Grammar course and were part of a six-week study abroad program in Salamanca, Spain. They were sixth-semester intermediate learners of Spanish and participated in the experiment for credit as one of their course requirements. The experiment took place on the second day of the study-abroad summer term. The average age of all learners was 19.66 (SD = 0.79, range = 18 to 21). There were 26 female and five male participants. Of these 31 participants, three female learners were subsequently excluded from the study because they had either (i) been raised bilingually (n = 2); or (ii) recently completed a study-abroad program lasting two or more months (n = 1).

3.3. Written instruments

3.3.1. Language History Questionnaire. Participants completed a Language History Questionnaire (Li, Zhang, Tsai, & Puls, 2013), which included basic demographic questions about their age, gender, and education, as well as more thorough questions about experiences with other languages. Learners completed this questionnaire during the first week of their stay abroad.

3.3.2. Spanish grammar proficiency test. We administered a 45-item grammar proficiency test (García-Amaya, 2012) to all participants. The test consisted of a short passage with a series of multiple-choice fill-in-the-blank options, which presented a diverse set of grammatical items. Participants received one point for each correct response, for a total of 45 points. Learners completed this test on the third day of the summer term. As will be seen in Section 3.7, the proficiency results were used as a control variable in our statistical models.

3.4. 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). However, given our focus on specific grammar rules, we devised new videos. For each target grammar structure, this included the process of generating original scripts, 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 grammar items, as well as their placement and randomization in each of the videos. The first and second authors used a Marantz Pmd620 digital recorder and two Shure WH20 head-mounted microphones to record the scripts.

The animated videos were created using Nawmal (www.nawmal.com), an animation program that allows users to create videos by choosing from a menu of predesigned characters and sets. This software allows for flexibility in the design, including the ability to upload user-recorded voices directly into the application — in our case, these were the recordings made by the two authors of this study — which are then automatically lip-synched to fictional characters. The Nawmal software also supports the inclusion of gestures as the characters go through their dialogue, as well as camera movements (e.g., close-ups, panning, dollying), which can help make the scenes feel more dynamic and natural.

Three unique animated videos were created, one per target grammar topic: *gustar*-type verbs, preterite/imperfect contrast, and the subjunctive in noun clauses (see Section 3.5). Within each video, there were three possible conditions for each target sentence: control NC (no-captioning) sentences that did not show any text for the target sentences; TE1 sentences that included target verbs highlighted in their entirety within the target sentences; and TE2 sentences, in which only the critical morphological and grammatical cues, and their relations, were highlighted. All target sentences were visually presented between square brackets in order to signal to participants that these sentences would need to be recalled. For the TE1 and TE2 conditions, 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. Table 3.1 offers a summary of the TE1 and TE2 manipulations per grammar topic.

We created three orders for each video so that the same target sentences would not appear in the same condition for all learners. For each of the three orders (for each video), the presentation of each of the three conditions was randomized (see Supplementary Materials; Appendix A, Table 3.A1). We adopted a within-subjects design, in which all participants saw all three

conditions for each of the three grammatical topics.

Table 3.1. Summary of the TE1 and TE2 Manipulations per Grammar Topic.

	Condition		
Grammar topic	Control	TE1	TE2
Gustar-type verbs	n/a	target verb is bold and yellow	target verb is bold and underlined; target morpheme is coded in yellow; all other plurality markers are coded in yellow, but not in bold
Preterite/ imperfect	n/a	target verb is bold and yellow	target verb is underlined; preterite target morpheme is bold and orange; imperfect target morpheme is bold and yellow
Subjunctive in noun clauses	n/a	main clause verb, conjunction *que* 'that', and subordinate subjunctive verb are in bold and yellow	main clause verb is bold and orange, followed by an orange bold arrow; conjunction *que* 'that' in bold and white; subjunctive verb is underlined, and target subjunctive morpheme is in bold and yellow

3.5. Grammar content

The specific grammar rules included in each video were taken from the course textbook *Repase y escriba: Curso avanzado de gramática y composición* (Canteli Dominicis & Reynolds, 2014).

Gustar–type verbs. The *gustar*-type verb construction is different from the English 'to like' construction: "despite their closeness in meaning, these predicates exhibit a divergent syntactic behavior: whereas 'like' codes as subject the entity that experiences a certain feeling, and as object the stimulus responsible for that feeling, *gustar* expresses the experiencer through an indirect object (or dative) and the stimulus through the subject" (Vázquez Rosa, 2006, p. 1).

For instance, in English, it is standard to construct a sentence that has the subject/experiencer "liking" a direct object (e.g., 'I like red roses'). However, in Spanish a different construction is used: *A mí me gustan las rosas* 'Red roses are pleasing to me', whereby the subject of the sentence is what in English would be considered the direct object, *rosas* 'roses'. The conjugated verb in the Spanish construction depends on whether the subject (i.e., *rosas*) is singular or plural. This L2 syntactic ordering poses challenges for L1 English speakers because it diverts from the canonical subject-verb-object (SVO) word order pattern found in English, instead favoring an OVS word order as its most frequent syntactic pattern (VanPatten et al., 2009). One common observation in L2 Spanish acquisition research is that, at the beginning stages, learners tend to interpret the subject as the first nominal feature in a sentence (Gass 1989; Lee and Malovrh, 2009; Seibert Hanson & Carlson, 2014; VanPatten, 1996). In the case of *gustar*-type verb structures, processing of the initial noun phrase or preverbal object pronoun as the subject can lead to a non-standard conjugation of the main-clause verb. However, there is evidence that target-like processing of OVS structures of this type can be promoted by instructional techniques in which the "connection between form and meaning is made virtually unequivocal and unavoidable" (DeKeyser & Prieto Botana, 2013, p. 456; see also Sanz & Morgan-Short, 2004; VanPatten & Oikkenon, 1996).

Contrary to the other target grammatical forms in the study, this structure does not involve the learning of a series of rules but rather understanding the non-canonical mapping of thematic roles in OVS sentences in order to correctly conjugate *gustar*-type verbs, as well as learning the particular lexical forms used in this construction.

In the current study 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 was presented four times in the animated video, twice in the singular form, and twice in the plural form.

Preterite/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 characterize past actions having a definitive beginning and endpoint (e.g., *caminé* 'I walked'), whereas imperfect forms characterize past actions or states being viewed as in progress (e.g., *caminaba* 'I was walking / I used to walk'). As noted in Liskin-Gasparro (2000), tense-aspect morphological forms such as the preterite and imperfect differ in their frequency distribution in the input received by L2 learners of Spanish, and thus lead to infrequent exposure of the contrast of these forms. However, as a motivating point for our study, Blyth (2005) asserts that such grammatical forms can benefit from interventions that "render surface forms more frequent and more salient, thereby allowing the learner to focus on form in a meaningful context" (p. 213).

For the preterite/imperfect, three rules for each form, and one rule which contrasted their usage were included in the animated video. Each rule was represented through four different verb instances within the video script. Given that the acquisition of these structures in L2 Spanish can be influenced by lexical aspect (Bardovi-Harlig, 2000)[9], our design controlled for this feature such that accomplishments and achievements were used in the preterite, and activities and states were used in the imperfect.

Subjunctive in noun clauses. The Spanish subjunctive mood is typically used in sentences with multiple clauses involving subordination, in which the subject of the main clause exerts

[9] The Lexical Aspect Hypothesis (Bardovi-Harlig, 2000) states that learners first use preterite morphology on achievements and accomplishments, later extending its use to activities and states, whereas they begin using imperfect morphology on states, extending next to activities, then to accomplishments, and finally to achievements.

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 learners of Spanish given its low frequency, and the low perceptual salience of the subjunctive inflection in the input (Collentine, 2013; DeKeyser & Prieto Botana, 2013). However, research has shown that breaking down the syntactic and inflectional components of this construction can facilitate its acquisition regardless of learners' readiness (Collentine, 2013; but see Leow et al., 2003). To this end, in the current study, both the verb in the main clause, which acts as a cue to the subjunctive, and the subordinated subjunctive verb, were textually enhanced in order to facilitate learners' understanding of the rules underlying the non-adjacent target forms (for further motivation, see discussion of Wilson et al. (2018), Section 2.1), either through TE1 or TE2.

For the subjunctive in noun clauses, five rules were targeted. Four different verb instances represented each rule.

3.6. Data collection

The experiment took place in a large auditorium where each participant was provided with a laptop and a headset. The full captioning experiment was programmed in OpenSesame (Mathôt, Schreij, & Theeuwes, 2012) and took participants less than one hour to complete. During the single experimental session, participants were presented with three animated videos on (1) *gustar*-type verbs, (2) preterite/imperfect forms, and (3) subjunctive in noun clauses. Learners saw each of the three videos in a random order. Each video included a pre-established number of sentences without TE (i.e., "no-captioning" or NC), others with TE1, and yet others with TE2, with the order of these conditions being randomized and controlled across three versions of each video (see Supplementary Materials; Appendix A, Table 3.A1).

Learners were informed that the majority of the videos would include Spanish-language captioning, which consisted of white, non-bolded 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, and once the subsequent audio of the spoken sentence had ended, the video paused for a maximum of 20 seconds. Learners knew that, during these 20 seconds, they had to type on their laptops verbatim all of the words that they could remember from the target sentence.

3.7. Data Analysis

Statistical analyses were conducted using R Studio version 1.0.143 (RStudio Team, 2015). The data were analyzed by multilevel generalized linear regression models utilizing the glmer() function within the lme4 package in R (Bates, Maechler, & Bolker, 2015).

For each trial, we measured the accuracy of the grammatical verbs from each target sentence. Each sentence received a score of 0 or 1 based on the correct usage of the target verb ending only. The dependent measures for each model were mean of trials correct.[10] Each model additionally included five fixed effects, two of which were predictor variables: GRAMMAR STRUCTURE (*gustar*-type verbs, the preterite/imperfect and subjunctive), and CAPTIONING (NC, TE1, and TE2); and three of which were control variables: GRAMMAR PROFICIENCY, SYLLABLES TO TARGET VERB (range 0-16) and SYLLABLES AFTER TARGET VERB (range 0-21).[11] All models included random intercepts for SUBJECTS.

[10] Only five trials were treated as missing for one participant whose program crashed near the end of the experiment. The trials corresponded to the *gustar*-type verbs video.

[11] The GRAMMAR PROFICIENCY measure was included as a fixed variable to take into account individual differences in Spanish proficiency. Given that the target sentences varied not only in length but also in terms of the position of the target verbs in the sentences, we also included SYLLABLES TO TARGET VERB and SYLLABLES AFTER TARGET VERB as fixed variables.

Since 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 GRAMMAR STRUCTURE and CAPTIONING.

4. Results

4.1. Accuracy

Figure 3.1 plots the mean accuracy score per target construction and TE manipulation (see Supplementary Materials; Appendix B; Table 3.B1 for a summary of the group accuracy proportion scores). In the following three sections, we offer a descriptive interpretation and statistical analysis of these findings.

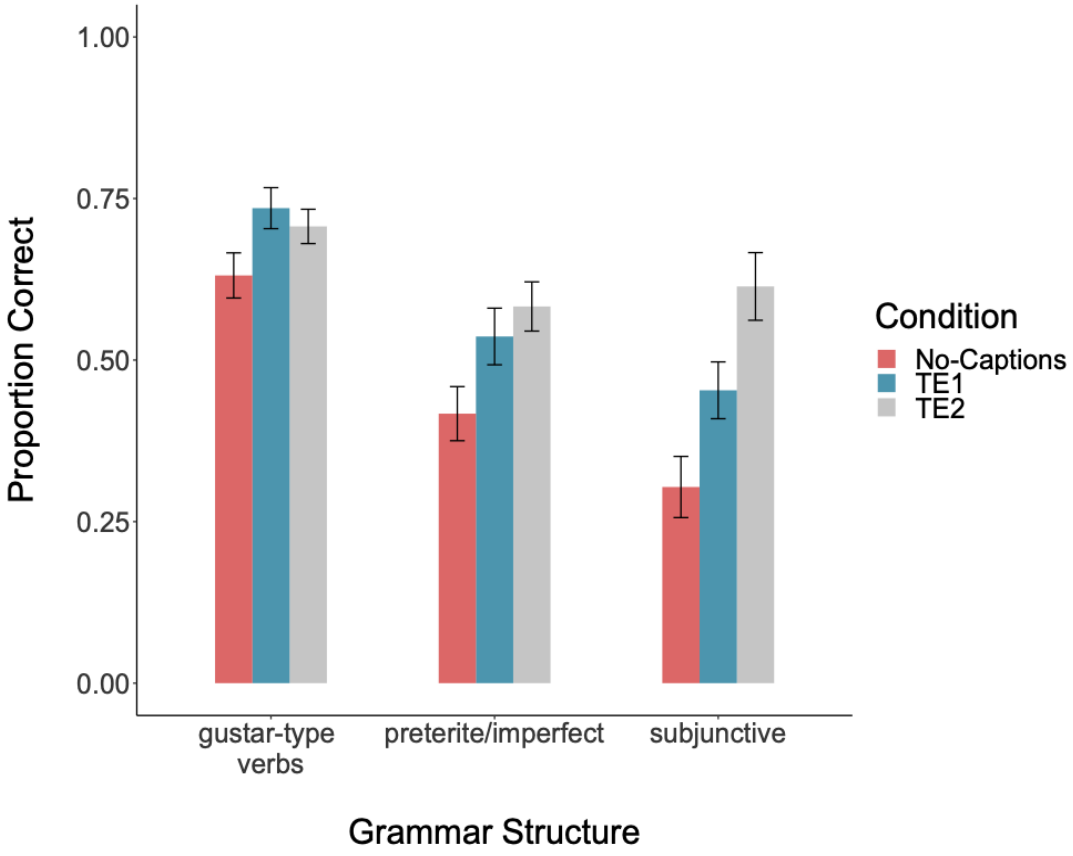

Figure 3.1. *Proportion correct scores for all grammar topics by condition. Error bars are 2 standard errors long.*

4.1.1. Gustar-type verbs

The *gustar*-type verbs data are plotted in the left-hand side of Figure 3.1. The pattern for this construction suggests an advantage of the two captioning conditions (i.e., TE1 and TE2) over NC.

To investigate the effects of captioning on accuracy, we ran a generalized linear mixed effects model which included fixed variables of GRAMMAR PROFICIENCY, SYLLABLES TO TARGET VERB, SYLLABLES AFTER TARGET VERB, as well as our predictor variables of interest: CAPTIONING and GRAMMAR STRUCTURE. The first model, with the *gustar*-type verbs construction and the NC condition as the reference levels, revealed a significant positive condition effect for TE1, β = 0.54, SE = 0.22, p = 0.01. The effect for TE2 was in the expected direction, but the model revealed only marginally significant results, β = 0.37, SE = 0.21, p = 0.08. In order to investigate if there were group differences between the two captioning conditions, we ran the same model with TE1 as the reference level. The model did not reveal a significant effect for TE2, β = - 0.16, SE = 0.22, p = 0.47. To summarize, the TE1 condition led to greater reproduction accuracy than the NC condition, and the TE2 condition showed marginal effects (over NC) in the expected direction. However, there was no significant difference between the two TE conditions.

4.1.2. The preterite/imperfect contrast

The middle panel of Figure 3.1 illustrates the preterite/imperfect accuracy data. As with the *gustar*-type verbs results, the data pattern suggests an advantage of captioning (TE1 and TE2) over the NC condition. The data dispersion additionally indicates a slight advantage for TE2 over TE1.

To investigate the specific effects of captioning on accuracy, we ran the same models as in the *gustar*-type construction but used the preterite/imperfect and the NC conditions as the reference levels for model 1, and the preterite/imperfect and the TE1 condition as the reference levels for model 2. Model 1 revealed significant positive effects for TE1, $\beta = 0.55$, SE = 0.19, $p = 0.003$, and for TE2, $\beta = 0.75$, SE = 0.19, $p < 0.001$, i.e., both captioned conditions led to greater reproduction accuracy than NC. Model 2, which focused on group differences between the two captioning conditions, did not reveal a significant effect for TE2, $\beta = 0.20$, SE = 0.19, $p = 0.28$.

4.1.3. The subjunctive in noun clauses

The results for the subjunctive in noun clauses are plotted on the right-hand side of Figure 3.1. Based on the data dispersion, there appears to be an advantage of the two captioning conditions over the NC condition. Additionally, and contrary to the patterns for the other grammar topics, the plots suggest an overall advantage for the TE2 condition.

Again, to investigate the effects of captioning on accuracy, we ran the same models as for the previous grammar structures but using the subjunctive and the NC conditions as the reference levels for model 1, and the subjunctive and the TE1 condition as the reference levels for model 2. Model 1 revealed significant positive effects for TE1, $\beta = 0.67$, SE = 0.23, $p = 0.003$, and TE2, $\beta = 1.42$, SE = 0.23, $p < 0.001$. Model 2, which investigated group differences between the two captioning conditions, revealed a significant positive effect for TE2, $\beta = 0.74$, SE = 0.22, $p < 0.001$, thus confirming our previous observation that there is an overall advantage for the TE2 manipulation.

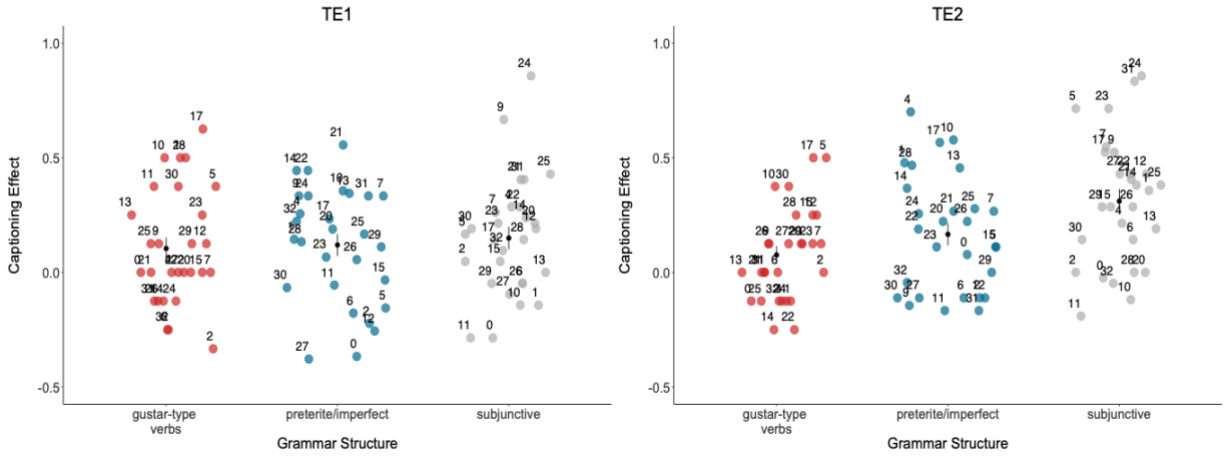

Figure 3.2. *By-subject TE captioning effects for each grammar topic. Error bars are 2 standard errors long*

4.1.4. Individual Data

In order to determine whether the effects uncovered in the previous sections are reliable across individual learners, we calculated individual captioning-effect scores for each participant per structure. This was done by subtracting each participant's average score for NC from their average TE1 and TE2 scores, respectively (see plots in Figure 3.2).

The data pattern illustrated in Figure 3.2 shows a relatively even distribution of individual scores across structures within each TE condition. Most, but not all, learners show some degree of sensitivity toward the captioning + TE manipulations (i.e., scores above 0). Nonetheless, in each instance, there are learners who show no effect or a negative effect of TE (i.e., scores at or below 0, respectively). Closer examination of the data spread also reveals that one learner, specifically Participant 2, is consistently among the lowest performers for all structures across both TE conditions: in all cases, Participant 2 shows either a negative effect of TE, no effect, or a

negligible positive effect (the six scores for this learner, in order of presentation from Figure 3.2, were: -0.333, -0.222, 0.048; 0.000, -0.111, and 0.000).

We would furthermore point that that, even when certain learners demonstrate a positive effect of captioning, it is not consistent across structures. For instance, Participant 24 showed greater sensitivity toward the TE1 and TE2 manipulations for the preterite/imperfect (scores = 0.333 and 0.255, respectively) and the subjunctive (scores = 0.857 and 0.857, respectively), but not for *gustar*-type verbs (scores = -0.125 and -0.125, respectively).

5. Discussion

In this study, we considered the effect of differential types of TE (or lack thereof) on L2 grammar uptake. Overall, we showed that captions incorporating some type of TE led to increased accuracy in learners' reproductions of the target grammatical forms relative to the non-captioned control conditions. This suggests that the provision of the TE led to greater attention and more optimal processing of target grammatical features. For two of the three target structures, namely, *gustar*-type verbs and the preterite/imperfect contrast, our analysis did not reveal significant differences between highlighting the full target verb (TE1) and the target morphemes that serve as cues to interpretation (TE2). On the other hand, for the subjunctive, providing learners with highlighting on both syntactic and inflectional cues (i.e., TE2) led to increased accuracy of TE2 over TE1. In Sections 5.1 through 5.3, we focus on specific effects for each structure and a more thorough discussion thereof.

5.1. Effects of TE on accurate uptake of gustar-type verbs

Regarding *gustar*-type verbs, this structure requires learners to understand the non-canonical mapping of thematic roles to properly conjugate verb forms using appropriate singular or plural morphemes. Learners must additionally learn the set of verbs that require this type of construction. In order to facilitate learner processing of this form, we either highlighted the target *gustar*-type verb in its entirety (TE1), or the target number morpheme of the verb, as well as all other number markers following this verb (TE2). The statistical results showed a significant effect of the TE1 over NC conditions and a marginally significant effect of TE2 over NC. These results provide additional support of the positive effects of TE on *gustar*-type verbs, both in L2 uptake (as revealed here) and in L2 learning, as demonstrated in Cintrón-Valentin et al. (under review; Chapter 2 of this dissertation). Whereas most previous literature on the acquisition of *gustar*-type verbs focuses on the processing and use of the clitic pronoun preceding the verb (e.g., Cerezo, Caras, & Leow, 2016; Lee & Malovrh, 2009), in our study we explored an additional, sometimes overlooked challenge in the acquisition of these structures, namely the agreement between verb morphology and its subject. We showed that by including TE in the multimodal media, learners can, in fact, overcome this acquisitional challenge during L2 uptake.

However, we did not uncover statistical differences between the two captioned conditions. One question that arises, based on this outcome, is why there was no added effect of TE2 relative to TE1. One possible explanation could be related to the nature of the syntactic dependencies in question. As discussed in Section 2.1 (recall Wilson et al., 2018), the learning of adjacent dependencies is relatively more straightforward than that of non-adjacent dependencies. In our design, the number markers following the target *gustar*-type verb morpheme were always adjacent to one other (e.g., *me molest<u>an</u> l<u>os</u> deport<u>es</u>*), except for three sentences (of the total set

of 24).[12] Our results, therefore, add to the TE literature by showing that in such constructions with adjacent dependencies, there may be no additional benefit in highlighting morphological cues to grammatical number.

5.2 Effects of TE on accurate uptake of the preterite/imperfect

Previous research on the effects of TE on the acquisition of Spanish past-tense aspect has tailored TE to either full lexical entries (e.g., Jourdenais et al., 1995; Leeman et al., 1995; Loewen & Inceoglu, 2018; Overstreet, 1998) or to aspectual morphemes of conjugated verbs only (e.g., LaBrozzi, 2016). To our knowledge, no single study has examined the possibility of differential effects of TE by comparing learner data from TE on a full lexical entry to data from TE on an inflectional morpheme. In our study, we compared the outcomes of NC, TE1, and TE2, and showed a significant positive effect of both TE1 and TE2 compared to NC, but not between TE conditions. In other words, and similar to *gustar-type* verbs, we did not find a significant difference between whole-word and inflectional morpheme TE on preterite/imperfect uptake.

The few studies examining the effects of TE on learner acquisition of the preterite/imperfect have yielded mixed findings, with some studies uncovering positive effects of TE on learners' noticing and production of these forms (e.g., Jourdenais et al., 1995; Leeman et al., 1995), whereas others have not (Cintrón-Valentín et al., under review (Chapter 2 of this dissertation); Overstreet, 1998). Overstreet discusses that the lack of a TE effect may be due to the added difficulty of learning how two forms function in contrast to one form within a specific semantic context. 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

[12] These were sentences that included the determiner *mucho* 'much' between the verb and subject.

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. In the case of the preterite/imperfect, at the local level, TE on these forms might actually distract learners' attention from the surrounding discourse, which offers critical information about the specific contexts in which each of the two aspectual choices are used (see also Bardovi-Harlig, 1998, regarding the importance of narrative context).

In studies that have shown positive effects of TE on learners' production of the preterite/imperfect forms (e.g., Leeman et al., 1995), learners in the enhancement condition may have benefited from an added compound enhancement (Han et al., 2008). Specifically, in Leeman et al. (1995), in addition to receiving TE combined with the provision of corrective feedback, learners received enhancement of forms inside and outside of the classroom, as well as the explicit instructions to focus on both meaning and form while processing the input. Additionally, as part of the TE, learners had the opportunity to attend to the linguistic forms at their discretion, and re-access previously presented text, thus allowing for more permanent visual substance of the textually enhanced forms, and more time to process the surrounding semantic context (see also Jourdenais et al., 1995).

In our study, in contrast, learners did not have the opportunity to re-access the previous discourse upon viewing an enhanced preterite or imperfect form — this methodological difference may help to explain the lack of significance between TE1 and TE2 in our results. Altogether, the collective findings on the preterite/imperfect open space for more nuanced TE designs where the benefits of presenting one form at a time (rather than two), as well as the need

for learners to re-access previous contextual information to achieve successful uptake, can be directly assessed.

5.3 Effects of TE on the uptake of the subjunctive in noun clauses

As mentioned previously, the Spanish subjunctive is a relatively complex morpho-syntactic structure emerging late in both L1 and L2 Spanish acquisition. Contrary to the other target structures in our study (see Section 6.1 and 6.2), the subjunctive is, in most instances in Spanish, restricted to subordinate clauses. Thus, for the L2 learner, the noticing and processing of this form in the input requires "a certain level of syntactic sophistication...such that processing beyond the matrix sentence can take place" (DeKeyser & Prieto Botana, 2013, p. 454). In our experiment, the verb in the main clause, which can act as a cue to the subjunctive, the subordinated subjunctive verb, as well as the relationship between the two cues, were made salient through the TE manipulations.

Our results showed significant effects of TE1 and TE2 over NC, as well as an added benefit of TE2 over TE1. The latter outcome contrasts with what we observed for *gustar*-type verbs and the preterite/imperfect. The additional positive effect of TE2 is in line with previous research on the learning of the subjunctive, which suggests that appropriate mood selection can be improved by providing learners with instructional strategies aimed at the optimal processing of these forms by breaking down a sentence's syntactic and inflectional components (e.g., Collentine, 2013). Our provision of TE2 for the subjunctive is likewise in line with the general tenets proposed in Wilson et al. (2018), whereby there is a successful effect of providing learners with visual cues to facilitate the grouping and processing of non-adjacent cues to interpretation.

5.4 Implications for research on TE

Until recently, the captioning research had primarily focused on its capacity to facilitate vocabulary learning and comprehension, with only one study investigating its potential in supporting L2 grammar learning through multimodal input (Cintrón-Valentín et al., under review, Chapter 2 of this dissertation). The results of Cintrón-Valentín et al. revealed that captioning + TE can, for some structures, aid in the acquisition of L2 grammar, particularly for *gustar*-type verbs and the subjunctive. Such mixed findings were in line with the research published in Han et al. (2008) and Lee and Huang (2008). Critically, the study by Cintrón-Valentín et al. investigated learner intake of the target forms but did not include a measure of immediate attention through an uptake design, like we created here. As mentioned previously, conscious attention to linguistic forms in the input is an essential step in L2 acquisition (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, but also their immediate uptake. We additionally contribute to this literature by showing that TE2 can offer an added positive effect on the uptake of non-adjacent dependencies (in the case of the subjunctive), but not necessarily on the uptake of adjacent dependencies (in the case of *gustar*-type verbs), as in Wilson et al. (2018). Future research may benefit from hypothesis-driven designs targeting the effect of TE in adjacent versus non-adjacent dependencies.

5.5 Implications for research on teaching L2 grammar

The current findings offer specific implications for pedagogical practices in the L2 classroom: Captions + TE can be a useful tool for L2 instructors. However, the optimal design of the TE manipulation – be it focused on a full lexical entry or the target morpheme and/or additional

sentential cues – should be carefully tailored to the target structure in question. For instance, in teaching non-adjacent structures such as the subjunctive, where the noticing and processing of the target form impose greater processing demands, learners might benefit from techniques that highlight the appropriate combination of morpho-syntactic considerations. One such example involves "if-clause" conditionals, where the type of conditional form used in the main clause is constrained by the condition established in the if-clause, thus creating a non-adjacent syntactic relationship. Rosa and Leow (2004) note that learners' abilities to extract patterns from specific instances of conditionals require higher levels of awareness, through explicit grammatical instruction regarding the relationship between its parts. Our subjunctive TE2 results bring to light the idea that captioned media can serve as an additional resource for exposing learners to complex morpho-syntactic structures through practical form-focused input, thereby elucidating grammatical relations in non-adjacent dependencies.

Regarding individual data patterns, previous research shows that the presentation of group tendencies may obscure distinct patterns of between-learner variation in language acquisition (e.g., Larsen-Freeman, 2018). Our results help to underscore this idea. For instance, we showed that one learner (Participant 2) displayed low captioning-effect scores across all structures and both TE manipulations: specifically, of the six individual captioning-effect scores, five of them were either at 0.000 or negative. As the literature on individual differences demonstrates, factors such as learner proficiency, attitude, motivation and modality preferences (i.e., visual, written, aural) can affect learners' receptiveness to different instructional interventions (see Dörnyei, 2005). As with any instructional method, one important takeaway from our study is that a single pedagogical technique will not be equally effective for all learners.

Altogether, we do not believe that TE alone will necessarily provide the most optimal means of instruction for all linguistic structures. As the literature on Form-Focused Instruction suggests, different grammatical forms require different levels of explicitness (Indrarathne & Kormos, 2017; Long, 2006; Spada & Tomita, 2010). Clearly, much work remains in terms of fine-tuning the quantity and types of enhancement needed for the successful acquisition of different constructions. This work should additionally consider the interaction of different configurations of TE with individual variables such as learners' L1, L2 proficiency, and prior knowledge, as demonstrated in Han et al. (2008).

5.6 Limitations and future directions

The study has its limitations, including (1) the lack of a non-enhanced captioned condition, (2) the inclusion of only one outcome measure, limited to the written modality, and (3) the short-term nature of the experiment.

5.6.1 The lack of a non-enhanced captioning condition. TE designs that include a direct comparison between enhanced versus unenhanced experimental conditions are more advantageous in directly addressing the unique contributions of TE in facilitating learner acquisition of the target grammatical forms (Leow & Martin, 2017). By providing this direct comparison, such designs are more equipped to tease apart whether the use of captioning is the single contributing factor to any positive effects in L2 learning. Given the design of our study, we were not able to tease apart any confounding effects of the written modality of captioning itself, from the incorporation of TE in addition to captioning. Future research could well incorporate a third manipulation, namely a captioning-without-enhancement condition, in order to address the concerns mentioned in Leow and Martin (2017).

5.6.2 The focus on one written outcome measure. Different treatments can render differential outcomes as a consequence of variations in test modality. For example, Sydorenko (2010), in a study on L2 vocabulary learning, presented English-speaking L2 learners of Russian with various experimental conditions that differed in the degree of aural and visual support included in the videos: Group 1 saw a video with captioned text and aural support; Group 2 saw a video with aural support only (i.e., no captions); Group 3 saw a video with captions only (i.e., no audio). All learners were subsequently tested in written and aural word recognition. Sydorenko found that Groups 1 and 3 scored higher on written than on aural recognition of word forms. Contrastingly, Group 2 scored higher in the aural word recognition tests. In our study, learners were offered aural support during all conditions, but written support in only two of the three conditions. We found significant positive effects of written support in a written assessment task but did not include aural assessment. A next logical step for future research could involve experiments designed to take into account the critical relation between input modality and test modality, focusing on grammar specifically.

5.6.3 The short-term nature of the experiment. The focus on learner uptake as a measure of learners' immediate recall of the material provides critical insight into learners' immediate noticing and processing of target forms but does not directly inform researchers about learners' subsequent learning of the attended form based on successful intake. Future studies including a battery of measures ranging in their implicitness/explicitness (Norris & Ortega, 2000), as well as long-term delayed testing of these measures, would allow us to develop a more complete understanding of grammar development through the usage of captioned media.

6. Conclusion

We have presented data from an innovative experiment designed to test the effect of TE video captions on learners' immediate uptake of three constructions in L2 Spanish (*gustar*-type verbs, preterite/imperfect, and the subjunctive). For all three structures, captions in addition to some form of TE led to increased accuracy in learners' uptake. This suggests that TE led to greater attention and more optimal processing of target grammatical features. We also uncovered an effect whereby form-focused TE2 offered an additive positive effect on L2 uptake for non-adjacent dependencies such as the subjunctive in noun clauses. Altogether, we have laid out a series of implications for L2 researchers and instructors, namely that: (i) incorporating some type of TE leads to increased accuracy in learners' L2 uptake; (ii) there are differential effects of TE based on the target structure; and (iii) TE does not yield uniform positive outcomes across learners, thus it should be viewed as one of multiple possible resources within the L2 classroom curriculum.

7. References

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259-302). Honolulu: University of Hawaii Press.

Allwright, R. L. (1984). The importance of interaction in classroom language learning. *Applied Linguistics, 5*, 156-171.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(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*(4), 299-320.

Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning, and use*. Oxford: Blackwell.

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

Blyth, C. (2005). From Empirical Findings to the Teaching of Aspectual Distinctions. In D. Ayoun & R. Salaberry (Eds.), *Tense and aspect in romance languages* (pp. 211-252). Amsterdam: John Benjamins.

Borrás, I., & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *The Modern Language Journal, 78*(1), 61-75.

Brown, R. (1973). *A first Language: The early stages*. Cambridge, MA: Harvard University Press.

Canteli Dominicis, M., & Reynolds, J. J. (2014). *Repase y escriba: Curso avanzado de gramática y composición* (7th ed.). St. John's University: New York.

Cerezo, L., Caras, A., & Leow, R. P. (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*(2), 26-291.

Cho, M. Y. (2010). The effects of input enhancement and written recall on noticing and acquisition. *Innovation in Language Learning and Teaching, 4*(1), 71-87.

Cintrón-Valentín, M. C, García-Amaya, L., Ellis, N. C. (under review). Captioning and Grammar Learning in the L2 Spanish classroom. *The Language Learning Journal.*

Cintrón-Valentín, M. C., & Ellis, N. C. (2016). Salience in second language acquisition: Physical form, learner attention, and instructional focus. *Frontiers in Psychology, 7, 1284.*

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*(1), 3-42.

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.

Comajoan, L. (2013). Defining and coding data: Narrative discourse grounding in L2 studies. In R. Salaberry, & L. Comajoan (Eds.), *Research design and methodology in studies on L2 tense and aspect* (pp. 309-356). Berlin: De Gruyter.

Comeaux, I., & McDonald, J. L. (2018). Determining the effectiveness of visual input enhancement across multiple linguistic cues. *Language Learning, 68*(1), 5-45.

DeKeyser, R. M. (1995). Learning second language grammar rules. *Studies in Second Language Acquisition, 17*(3), 379-410.

DeKeyser, R. M. (2005). What makes learning second language grammar difficult? A review of issues. *Language Learning, 55*(1), 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.

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. New York: Cambridge University Press.

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*(2), 164–94.

Ellis, N. C. (2017). Salience in usage-based SLA. In S. Gass, P. Spinner, & J. Behney (Eds), *Salience in second language acquisition (*pp. 21-39*)*. New York: Routledge.

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, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning, 51,* 281-326.

Gass, S. M., Spinner, P., & Behney, J. (Eds.). (2017). *Salience in second language acquisition*. New York: Routledge.

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.

Gaillard, S., & Tremblay, A. (2016). Linguistic proficiency assessment in second language acquisition research: The elicited imitation task. *Language Learning, 66*(2), 419-447.

Garza, T. J. (1991). Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals, 24*(3), 239-258.

Gass, S. M., Spinner, P., & Behney, J. (Eds.). (2017). *Salience in second language acquisition*. New York: Routledge.

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(*2), 373-402.

Han, Z., Park, E. S., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics, 29*(4), 597-618.

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period of second language acquisition: Evidence from 2/3 million English speakers. *Cognition, 177*, 263-277.

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*(1), 75-96.

Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions. *Studies in Second Language Acquisition, 38*(3), 401-430.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*(4), 541-577.

Jourdenais, R. M., Ota, S. Stauffer, B. Boyson, and 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: University of Hawaii Press.

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.

LaBrozzi, R. M. (2016). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research, 20*(1), 75-91.

Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning, 26*, 125-134.

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, B., & 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. *The Modern Language Journal, 102*(3), 557-577.

Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning, 57*(1), 87.

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*(3), 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* (Tech. Rep. No. 9, pp. 217–258). Honolulu: University of Hawai‘i, Second Language Teaching and Curriculum Center.

Leow, R. P. (1997). The effects of input enhancement and text length on adult L2 readers' comprehension and intake in second language acquisition. *Applied Language Learning, 8*, 151-182.

Leow, R. P. (2001). Do learners notice enhanced forms while interacting with the L2? An online and offline study of the role of written input enhancement in L2 reading. *Hispania, 84*, 496-509.

Leow, R. P., Egi, T., Nuevo, A. M., & Tsai, Y. C. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning, 13*, 93-108.

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*(03), 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.

Loewen, S., & Inceoglu, S. (2016). The effectiveness of visual input enhancement on the noticing and L2 development of the Spanish past tense. *Studies in Second Language Learning and Teaching*, *6*(1), 89-110.

Long, M. H. (2006). *Problems in SLA*. Mahwah, NJ: Erlbaum.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*, 314-324.

Markham, P. L. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals, 32*(3), 321-328.

Markham, P. L. (1993). Captioned television videotapes: Effects of visual support on second language comprehension. *Journal of Educational Technology Systems, 21*(3), 183-191.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41*(3), 720-739.

Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology, 18*(1), 118-141.

Muñoz, C. (2017). The role of age and proficiency in subtitle reading. An eye-tracking study. *System, 67*, 77-86.

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

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*(2), 127-162.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis.*Language Learning, 50*(3), 417-528.

Overstreet, M. (1998). Text enhancement and content familiarity: The focus of learner attention. *Spanish Applied Linguistics, 2*, 229-258.

Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 34(1), 80.

Price, K. (1983). Closed-captioned TV: An untapped resource. *MATESOL Newsletter, 12*(2), 1-8.

Rosa, E., &, Leow, R. (2004). Computerized task-based exposure, explicitness, type of feedback, and Spanish L2 development. *The Modern Language Journal, 88*(2), 192-216.

RStudio Team (2015). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA. URL: http://www.rstudio.com/.

Sanz, C., and Morgan-Short, K. (2004). Positive evidence versus explicit rule presentation and explicit negative feedback: A computer-assisted study. *Language Learning, 54*(1), 35-78.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Seibert Hanson, A., & Carlson, M. T. (2014). Global effects: The roles of first language and proficiency in L2 processing of Spanish clitics. *Language Learning, 64*(2), 310-342.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15,* 165-179.

Shook, D. J. (1994). FL/L2 reading, grammatical information, and the input to intake phenomenon. *Applied Language Learning, 5*, 57-93.

Slimani, A. (1992). Evaluation of classroom interaction. In C. Anderson, & A. Beretta (Eds.), *Evaluating second language education* (pp. 197-221). Cambridge: Cambridge University Press.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*(2), 263-308.

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology, 14*(2), 50-73.

Tolentino, L. C., & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning, 64*(2), 279-309.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*(2), 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.

Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language teaching, 43*(01), 1-37.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood, NJ: Ablex.

VanPatten, B. (2003). *From input to output: A teacher's guide to second language acquisition*. New York: McGraw-Hill.

VanPatten, B., & Oikkenon, S. (1996). Explanation vs. structured input in processing instruction. *Studies in Second Language Acquisition, 18*, 495-510.

VanPatten, B., Inclezan, D., Salazar, H., & Farley, A. P. (2009). Processing instruction and dictogloss: a study on object pronouns and word order in Spanish. *Foreign Language Annals, 42*(3), 557-575.

Vázquez Rozas, V. (2006). Gustar-type verbs. In J. Clancy 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*(1), 54-73.

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford: Oxford University Press.

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,* 1-16.

Winke, P. M. (2013). The effects of input enhancement on grammar learning and comprehension. *Studies in Second Language Acquisition, 35*(2), 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*(1), 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.

Wong, W. (2003). Textual enhancement and simplified input effects on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning, 13*(2), 109-132.

Wulff, S., & Ellis, N. C. (2018). Usage-based approaches to SLA. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 37-56). Amsterdam: John Benjamins.

Appendix A.

Design

Table 3.A1. Item randomization by target structure, order and condition.

Structure	**Number of Target Sentences**	**Order 1**			**Order 2**			**Order 3**		
		Pret./ Imp	*Gustar*	Subj.	Subj.	Pret./ Imp.	*Gustar*	*Gustar*	Subj.	Pret./ Imp.
		NC	**TE1**	**TE2**	**NC**	**TE1**	**TE2**	**NC**	**TE1**	**TE2**
Gustar	24	8	8	8	8	8	8	8	8	8
Pret./Imp.	28	9	9	10	9	10	9	10	9	9
Subjunctive	20	6	7	7	7	6	7	7	7	6

Appendix B.

Additional data tables

Table 3.B1. Accuracy Proportion Scores for each grammar topic by condition.

Condition	Mean	SD	95% CI
	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		
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. CI = confidence interval

CHAPTER 4. The Effects of Form Focused Instruction and Captioned Media in L2 Development: A Follow-up Study

1. Introduction

1.1. Overview

There is increasing interest in the use of multimedia learning techniques, such as captioned media, as a means to promote second language (L2) comprehension, vocabulary learning (e.g., Vanderplank, 2010; Montero-Perez, Van den Noortgate & Desmet , 2013) and more recently, grammar development (e.g., Cintrón-Valentín, García-Amaya & Ellis, under review; Chapter 3 of this dissertation; Lee & Révész, 2018). Guided by theoretical principles from the Second Language Acquisition (SLA) and attention literatures (e.g., Doughty & Williams, 1998; Schmidt, 2001; Sharwood Smith, 1993), the latter studies have found positive effects of visually enhanced captions on various aspects of grammar development (e.g., receptive knowledge, written production, learner uptake). In an effort to better understand the role of captioned media on L2 grammar development, we offer a direct follow-up to Cintrón-Valentín et al.(under review) and Chapter 3 of this dissertation. Specifically, (i) we examine the effect of Textual Enhancement (Sharwood Smith, 1993; henceforth TE) in combination with captioned video on the L2 acquisition of various elements of vocabulary and Spanish morphosyntax (i.e., the preterite/imperfect contrast, *gustar-type* verbs, the subjunctive in noun clauses, and the conditional); and (ii) we explore the effects of TE captioned video in combination with, and in the absence of, explicit grammar instruction on learner knowledge of the targeted structures.

2. Background

2.1. SLA with multimedia and captioning

SLA with multimedia can be defined as "the use of words and pictures [either static or dynamic] to provide meaningful input, facilitate meaningful interaction with the target language, and elicit meaningful output" (Plass & Jones, 2005, p.469). Elaborating on such instructional methods, Webb and Nation (2017) discuss how the use of elaboration techniques, designed to enrich a learner's knowledge of a word "by encountering more aspects of its form, meaning, and use", such as the inclusion of pictures in addition to written text, can in many instances ,"provide a memorable image of the meaning and context of a word", p.73), and thus facilitate acquisition. Captioned media, where the text is presented in the same language as the audio (Jung, 1990), can be considered one of many multimedia materials available to L2 learners and instructors (see for instance Chun & Plass, 1996, 1997; Jones and Plass, 2002, for additional examples). This technique has garnered increasing attention in recent years given its demonstrated benefits in facilitating L2 comprehension and vocabulary acquisition (e.g, Vanderplank, 2010; Montero-Perez, et al., 2013). On the benefits of captions for L2 vocabulary learning, for instance, 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 & Lafayette, 1994, p.70).

More specifically, 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 & Behney, 2017; Schmidt, 2001; Tomlin & 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. Vanderplank's (2016) model of language acquisition through captioned media similarly emphasizes 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.

2.2. Perceptual Salience, Form Focused Instruction and captioning

The potential role of captioned media in mediating learner attention to linguistic forms in the input is particularly relevant to L2 grammar development. Specifically, research on the role of attention in L2 acquisition suggests that low perceptual salience of grammatical forms in the input is largely responsible for learner challenges in grammar acquisition (e.g., Ellis, 2017; Gass, Spinner & Behney, 2017; Goldschneider & DeKeyser, 2001; Larsen-Freeman, 1976). For instance, there is evidence that the low perceptual salience of certain grammatical features, such as inflectional suffixes that vary based on the tense/mood/aspect of the grammatical context, contributes to L2 learners' difficulty in acquiring them (Cintrón-Valentín & Ellis, 2016; Ellis, 2006; Gass et al., 2017; Goldschneider & DeKeyser, 2001). In the L2 and grammar acquisition literature, one way of counteracting the effects of low-salience grammatical forms is by providing learners with Form Focused Instruction (FFI) techniques designed to render target structures more salient. FFI encapsulates a wide range of instructional activities that look to draw

learners' attention to linguistic forms in the input that might otherwise be ignored (Spada, 1997; Spada and Tomita, 2010; Ellis, 2012).

Two FFI methods that have been widely investigated both in SLA research and practice (see for instance: Cintrón-Valentin & Ellis, 2015, 2016; Han, Park, & Combs, 2008; Lee & Huang, 2008; Spada & Tomita, 2010; Norris & Ortega, 2000) are explicit grammar instruction (EGI) and Textual Enhancement (TE). Terrell (1991, p. 53) defines EGI as "the use of instructional strategies to draw the students' attention to, or focus on, form and/or structure," with instruction targeted at increasing the salience of inflections and other commonly ignored features by, first, pointing them out and explaining their structure and, second, providing meaningful input that contains many instances of the same grammatical meaning-form relationship. TE on the other hand, uses visual manipulations such as color-coding, boldfacing and underlining, providing a more unobtrusive means of increasing learners' awareness of non-salient forms in the input (Sharwood-Smith, 1993; Doughty and Williams, 1998). Given the increased reliance on multimedia materials in L2 teaching and learning, it is of growing interest to investigate how traditional instructional techniques such as FFI can be effectively integrated with such multimedia approaches to language learning.

Recent studies in the captioning and vocabulary learning literature have begun to implement similar techniques (e.g., Pujadas & Muñoz, submitted; Montero Perez et al., 2013). Pujadas and Muñoz (submitted), for instance, investigated the role of captioned video and focused instruction (i.e., being instructed on the target vocabulary words prior to the presentation of the captioned video) on the learning of vocabulary, revealing significant learning effects for learners' in the focused condition. 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 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). Overall, the findings of these studies suggest, that explicit instructional techniques in combination with captioned media and FFI can make vocabulary more salient for learners and promote the learning of form-meaning connections.

To our knowledge, only three studies (Chapter 3 of this dissertation; Cintrón-Valentín, García-Amaya & Ellis, under review; Lee & Révész, 2018) have investigated the role of FFI in combination with captioned media on enhancing learner attention to, and optimizing the learning of grammatical forms. Lee and Révész (2018), for instance, investigated the effects of TE-captioned media on the learning of pronominal anaphoric reference in L1 Korean learners of English through a series of multimodal input-based activities finding significant effects of captioning on learners' receptive grammar knowledge. This study, however, did not directly investigate captioned videos nor did they provide learners with pictures aimed at guiding the narrative presented through the bimodal input (aural and written).

Cintrón-Valentín et al. (under review, see Chapter 2 of this dissertation) investigated how captioned video could serve as a useful tool for advancing grammar learning in the L2 Spanish classroom. Their study focused on four Spanish grammar structures, showing significant effects of TE-captions on some, but not all target forms, namely that of *gustar*-type verbs and the

subjunctive in noun clauses. However, several methodological limitations impacted the interpretability of their findings: (1) they did not consider the relative influence of different types of TE on grammar and vocabulary learning, and whether different types of TE have differential effects on L2 learning (see for instance, LaBrozzi, 2016); (2) they did not include a pretest prior to conducting the experimental sessions, making it difficult to tease apart any possible confound regarding the gains acquired through the treatment from pre-existing knowledge; (3) all captioned videos were fronted by an explicit grammar lesson making it difficult to tease apart whether the use of captioning was the single contributing factor to any positive effects in the learning assessments.

Chapter 3 of this dissertation addresses the first concern, as it pertains to grammar learning, by assessing effects of different designs of TE video captions on learners' immediate uptake of three grammatical constructions in L2 Spanish (*gustar*-type verbs, the preterite/imperfect contrast, and the subjunctive). Critically, whereas Cintrón-Valentín et al. (under review; Chapter 2) only included one type of TE-captions on grammar, highlighting the full lexical entry (e.g., the complete verb form containing the target morpheme), Chapter 3 of this dissertation included two types of TE-captions: a TE1-condition which presented target verbs highlighted in their entirety; and a TE2-condition which presented highlighting of the critical morphological and grammatical cues, and their relations. Their results showed that captions incorporating some type of TE led to increased accuracy in learners' reproductions of the target grammatical forms relative to the non-captioned control conditions. For two of the three target structures, namely, *gustar*-type verbs and the preterite/imperfect contrast, the analyses did not reveal significant differences between highlighting the full target verb (TE1) and the target morphemes that serve as cues to interpretation (TE2). On the other hand, for the

subjunctive, providing learners with highlighting on both syntactic and inflectional cues (i.e., TE2) led to increased accuracy of TE2 over TE1. In the current study we incorporate the TE2 manipulations included in Chapter 3 of this dissertation, specifically tailoring TE on the target morpheme only, along with any relevant grammatical dependencies. In order to address limitation 2, we additionally include a pretest of the targeted grammar forms in order to measure any potential learning gains following the captions + TE intervention and thus discern any effects of prior knowledge from the experimental treatment.

As a means of addressing limitation 3, the current study includes an experimental Grammar group, which did not receive explicit instruction prior to the animated captioned video. Critically, we do not believe TE-captions alone will necessarily provide the most optimal means of instruction for all linguistic structures. As the literature on Form-Focused Instruction suggests, different grammatical forms might require different levels of explicitness and explanation (Long, 2006, chap. 5; Spada & Tomita, 2010; Tolentino & Tokowicz, 2014). For instance, recent research by Indrarathne and Kormos, (2017) emphasizes the relevance of this important consideration. Indrarathne and Kormos (2017) investigated the effects of TE on the learning gains of Sri Lankan learners of English. Four experimental groups received different types of input texts where the grammatical target was presented in one of four conditions: unenhanced; enhanced; enhanced plus instruction; and enhanced plus instruction plus explanation. Participants' eye-movements while reading the texts were tracked, and the amount of attention to target grammatical items was measured. Two pre-/post-tests (Sentence reconstruction and Grammaticality Judgement) were used to assess learning gains. Overall, the results showed: (i) more learning in the explicit conditions, and (ii) a clear association between learning gains and attentional processing. The study demonstrates that if learners are given limited support in what

to pay attention to in the input, despite abundant examples in the text or even if they are visually enhanced, learners' attentional processes may nevertheless still not be directed to the target feature.

The key methodological differences between Cintrón-Valentín et al. (under review; Chapter 2) and the current study should allow us to better assess whether captioned videos can be effective in improving learner knowledge of L2 grammar within the L2 Spanish classroom setting, and more specifically, if there are differential effects based on the grammatical structures in question.

2.3. Research Questions

The current study aimed to extend previous research on captioning and second language acquisition. The study had five specific aims:

(1) to examine the effects of full captions + TE vocabulary on improving learner knowledge of vocabulary

(2) to investigate if any initial gains of full captions + TE vocabulary on the production of vocabulary are maintained over time.

(3) to examine the effects of full captions + TE grammar on improving learner knowledge of grammar

(4) to investigate if any initial gains of full captions + TE grammar on the production of grammar are maintained over time.

(5) To investigate if the effects of full captions + TE grammar are equally facilitative in the absence of explicit instruction.

    a. to examine the effects of full captions + TE grammar on improving learner

knowledge of grammar

b. to investigate if any initial gains of full captions + TE grammar on the production of grammar are maintained over time.

Similar to Cintrón-Valentín et al. (under review; Chapter 2 of this dissertation), 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 can 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. Contrary to Cintrón-Valentín et al. (under review; Chapter 2 of this dissertation), we additionally include RQ2 in order to investigate retention of the learned vocabulary words following the lab session. As indicated by Montero Perez et al. (2013), experimental designs considering the long-term effects of captioning on vocabulary retention through delayed posttests are scarce but necessary in order to put together a more complete picture of vocabulary development through the usage of captioned media.

4. Method

4.1. Participants

A total of 369 English-speaking L2 learners of Spanish were recruited from a Spanish Grammar course at a large Midwestern University. They were fifth-semester intermediate learners of Spanish and participated in the study for credit as part of one of their course requirements[13]. The course had 21 sections, which were quasi-randomly assigned to one of four

[13] Participants were fifth semester learners of Spanish or had received a high score in their Advanced Placement Spanish course in high school.

groups: a Lesson + No Salience group (Lesson + Control); a Lesson + Salience on Vocabulary (Lesson +SV) group; a Lesson + Salience on Grammar (Lesson +SG) group, and a No Lesson + Salience on the grammatical features group (No Lesson + SG) (see Table 4.1 for descriptive statistics). Of these participants, 63 (Lesson + Control = 18; Lesson+ SV = 14; Lesson +SG = 9; No Lesson + SG= 22) were excluded from the study (1) if they had and L1 other than English or early experiences with other languages (n = 35); (2) if they had been exposed to the Spanish language before age 6 (n = 9); or (3) if they had participated in a L2 Spanish study-abroad experience for two months or more (n = 19).

Table 4.1. Descriptive Statistics.

Group	**N subjects**	**Age Range**		**Mean Age (SD)**	**Sex**	
		Minimum	Maximum		Females	Males
Lesson + Control	89	17	35	18.74 (1.96)	58	27
Lesson + SV	84	17	28	18.74 (1.45)	53	27
Lesson + SG	88	17	29	18.59 (1.65)	63	23
No Lesson + SG	108	17	24	18.44 (1.10)	70	31

Note. These were several participants who did not report their sex (Lesson + Control = 4; Lesson + SV = 4; Lesson + SG = 2; No Lesson + SG = 7)

4.2. Target Structures

The target structures in the current study follow that of Cintrón-Valentín et al. (under review; Chapter 2 of this dissertation) with two exceptions: we did not investigate the *ser* and *estar* contrast, including instead the conditional tense.

Session 1.The preterite/imperfect. The standard usage of the two simple past tenses in

Spanish requires that learners understand the aspectual distinction between the preterite and imperfect forms (Colomé, 2013). Preterite forms are characterized as actions having a definitive beginning and endpoint (e.g., caminé 'I walked'), whereas imperfect forms indicate past actions or states being viewed as in progress (e.g., caminaba 'I was walking / I used to walk'). As noted in Liskin-Gasparro (2000), tense-aspect morphological forms such as the preterite/imperfect differ in their frequency distribution in the input received by L2 learners of Spanish, and thus lead to infrequent exposure of the contrast of these forms. Blyth (2005) asserts that such grammatical forms can benefit from interventions that "render surface forms more frequent and more salient, thereby allowing the learner to focus on form in a meaningful context" (p. 213).

Session 2. Gustar-type verbs. L1 learners' mastery of the gustar-type verb construction is considered especially challenging given the marked differences between its English counterpart 'to like' *(e.g., Cerezo, Caras & Leow, 2016):*

> "Despite their closeness in meaning, these predicates exhibit a divergent syntactic behavior: whereas 'like' codes as subject the entity that experiences a certain feeling, and as object the stimulus responsible for that feeling, gustar expresses the experiencer though an indirect object (or dative) and the stimulus through the subject" (Vázquez Rosa, 2006, p. 1).

For instance, in English, it is standard to construct a sentence that has the subject/experiencer "liking" a direct object (e.g., 'I like red roses'). However, in Spanish a different construction is used: *A mí me gustan las rosas* 'Red roses are pleasing to me', whereby the subject of the sentence is what in English would be considered the direct object, *rosas* 'roses'. The conjugated verb in the Spanish construction depends on whether the subject (i.e., *rosas*) is singular or plural. This L2 syntactic ordering poses challenges for L1 English speakers because it diverts from the

canonical subject-verb-object (SVO) word order pattern found in English, instead favoring an OVS word order as its most frequent syntactic pattern (VanPatten et al., 2009). One common observation in L2 Spanish acquisition research is that, at beginning stages, learners tend to interpret the subject as the first nominal feature in a sentence (e.g., Lee and Malovrh, 2009; Seibert Hanson & Carlson, 2014; VanPatten, 1996). In the case of *gustar*-type verb structures, processing of the initial noun phrase or preverbal object pronoun as the subject can lead to a non-standard conjugation of the main-clause verb. However, there is evidence that target-like processing of OVS structures of this type can be promoted by instructional techniques in which the "connection between form and meaning is made virtually unequivocal and unavoidable" (DeKeyser & Prieto Botana, 2013, p. 456; see also Sanz & Morgan-Short, 2004; VanPatten & Oikkenon, 1996). Whereas most 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 & 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.

Session 3. 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, p.454; Collentine, 2012). 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, 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.

Session 4. The conditional tense. Conditional sentences are considered to be highly complex structures in both L1 and L2 acquisition, not only because of their morphosyntactic complexity, but also given the semantic complexity involved in learners' processing of this form (e.g., López Ornat, 1994). Given the generalized difficulty in acquiring this structure, the conditional is usually not taught in the first two years of language study. In general terms, the Spanish conditional tense is used to express probability or hypotheses about the past, present or future (Areizaga Orube, 2009). In the present study, we focus on one specific usage of the conditional, the expression of speculation or probability about the past using the 'must have + verb construction' (e.g., Where was John last night? He wasn't at home. He **must have been** in the lab/¿Dónde estaba John anoche? No estaba en casa**. Estaría** en el lab). We targeted a very specific low frequency usage of this construction, deviating from the usage included in the course textbook and in the learners' regular class discussion. In doing so, we aimed to explore how facilitative TE-captioned media would be in improving learner knowledge of a structure for which learners are known to have minimal experience with. Learners' difficulty with, and minimal knowledge of the conditional construction we targeted here was confirmed by testing the same learners who participated in Study 2 (Chapter 3) on their ability to produce this structure. With the exception of one learner, who had been taught how to use the conditional tense in High School, all students performed below chance, and reported not having learned this structure either in High School or in their College grammar courses.

4.3. Written instruments

4.3.1. Language History Questionnaire. Participants completed a Language History Questionnaire (Li, Zhang, Tsai, & Puls, 2013), which included basic demographic questions about their age, gender, and education, and more thorough questions about their experience with different languages.

4.3.2. Spanish vocabulary proficiency test. The Lextale-ESP (Izura, Cuetos & 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 recognized as Spanish words. As recommended by Lemhöfer & Broersma (2012) and Brysbaert (2013), the test was scored using the following formula:

$$\text{Score} = \text{N ‘yes’ to words} - 2 * \text{N ‘yes’ to nonwords.}$$

This scoring formula penalizes for guessing behavior, so that a participant who marks all words and nonwords as known, or one who answers randomly, would receive a score of 0.

In order to control for any possible familiarity of the target vocabulary items we additionally included the target vocabulary words and foils in this test. The target vocabulary words were coded and scored separately. Participants received one point for each target vocabulary word they recognized as Spanish, for a total of 23 points.

4.3.3. Elicited Imitation Task. Participants completed an Elicited Imitation (EI) task in order to measure their global Spanish proficiency. This task was originally developed by Ortega, Iwashita, Rabie and Norris (1999). Here we used the revised version included in Bowden (2016). This technique requires that the learner listen to a sentence and repeat it as exactly as possible

with the assumption that learners should be more successful in repeating sentences that contain semantic and morphosyntactic features that are familiar or known to them, and less successful in those where the linguistic structures exceed their knowledge. During the task, participants were aurally presented with 30 sentences, of increasing length (7-17 syllables), one at a time. Directly after the presentation of each sentence there would be a two second pause, followed by a 0.5 second tone sound that would serve as a cue for participants to repeat as much of the sentence as they could recall. The scoring criteria used for this study was the same as in Bowden (2016) and Ortega et al., (1999) which focused on the number and quality of the idea units produced by the speakers. Specifically, participants' utterances were scored on a 0-4 scale – a minimum score of 0 was given to instances of silence, unintelligible productions or minimal repetitions; and a maximum score of four points was given to exact repetitions. To ensure reliability of these scores, each individual test was independently scored by two raters, and any discrepancies in their scores were resolved prior to analyzing the data.

4.3.4. *Grammar Pretest.* Participants additionally completed a grammar pretest which included a representative sample of each of the target structures they would be tested on during the experiment. The test consisted of 51 production items, where the speakers were either asked to translate target verbs from English to Spanish or provide the correct Spanish usage of certain verbs based on specific sentential contexts (see Supplementary Materials; Appendix A).

4.3.5. Immediate posttests

4.3.5.1. Vocabulary recognition test. Participants were tested on their recognition of target vocabulary (see Table 4.2). 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 23 target words were tested as well as an additional 23

foils. A score of 1 was given for each correctly identified target word.

Table 4.2. 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	-
dormilonas	2	noun	-
caniches	2	noun	0. 36
sobremesa	2	noun	0. 71
impúdico	2	adjective	0. 89
espejuelos	2	noun	-
holgazán	2	adjective	0. 18
estantería	3	noun	2. 66
vergel	3	noun	1. 07
alambrado	3	noun	0. 36
boceto	3	noun	1. 07
valija	3	noun	2. 31
atolondrado	3	adjective	0. 18
alhajas	4	noun	1.78
antro	4	noun	1.78
huésped	4	noun	7.46
lingotazo	4	noun	0.18

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

4.3.5.2. Vocabulary translation test. 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 or other related words not presented in the movie were scored as incorrect.

4.3.5.3. Grammar recall of form task. Participants were tested on their ability to recall target grammatical forms. They were presented with multiple sentences and were instructed to fill in the blank with the correct conjugated verb. The responses were scored based on 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. Participants received a score of 0 for all incorrect inflections, and a score of 1 for all correct inflections.

4.3.5.4. 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 following the same procedure as for the recall of form task.

4.3.6. Two-week delayed posttests. Approximately two weeks after each of the four experimental sessions, similar versions of the grammar and vocabulary translation tests were administered during learners' regular class time in order to measure retention over time. For the grammar portion, the tests included the same verb items the learners had been tested on in the immediate posttests, but in different sentence contexts. The scoring for this test was the same as for the immediate vocabulary and grammar translation tests.

4.4. Grammar Lesson Videos

For each grammatical structure, a short grammar video lesson was created. Each video lesson

summarized 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. These were the same grammar lessons presented in Chapter 2 of this dissertation. In each practice exercise, participants were presented with a question on the target structure. They were given ten seconds to work through the question on their own and subsequently were provided with the correct answer.

4.5. 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 and rules, 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 randomization in each of the videos.

The animated videos were created using Nawmal (www.nawmal.com), an animation program that allows users to create videos by choosing from a menu of predesigned characters and sets. This software allows for a great degree of flexibility in the design, including the ability to upload user-recorded voices directly into the application, that is then automatically lip-synched to fictional characters. The Nawmal software also supports the inclusion of gestures as the characters go through their dialogue, as well as camera movements (e.g., close-ups, panning, dollying), 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 (No Captions, Salience on Vocabulary or Salience on Grammar). For each video, captions were added using SRT Edit Pro (http://www.finalsub.com/sep.html), which allowed for the inclusion of color-coding and bold-facing within the captioning line.

4.5.1. Vocabulary Content

The animated videos created for each lab session included target vocabulary–overall a total of 23 target words were included in the experiment (see Table 4.2 for the breakdown of these target words by session). The target vocabulary chosen for the experiment were either low-frequency words taken from the NIM Frequency database[14] (Guash, Boada, Ferré & Sánchez - Casas, 2013), or regional vocabulary words to which participants would have only been exposed if they 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. The videos for the preterite/imperfect,, gustar-type verbs and the subjunctive in noun clauses included in Cintrón-Valentín, García-Amaya and Ellis (under review; Chapter 2 of this dissertation) all followed a similar design. For each of these videos there were as many unique target vocabulary 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 vocabulary 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).

[14]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.

For the conditional video, we included four vocabulary items. Each vocabulary item was repeated twice, and contrary to the other three structures, the vocabulary items were spread out across the script rather than massed.

4.5.2. Grammar content

The specific grammar rules included in the original Cintrón-Valentín, García-Amaya and Ellis (under review; Chapter 2 of this dissertation) videos were taken from the course textbook *Repase y escriba: Curso avanzado de gramática y composición* (Cantelis Dominicis & Reynolds, 2014). For the conditional video, we focused on the expression of speculation or probability about the past. a special case which departed from the specific examples included in their course syllabus and textbook. These instances were selected by the research team and were piloted during the previous semester.

Depending on the target structure, either two or three rules, and one of their corresponding verb instances were included in the grammar lesson video. These same items, as well as the remaining rules and verb instances, appeared in the animated video.

Session 1: Preterite and Imperfect. For the preterite/imperfect, three rules for each simple past form, and one rule which contrasted their usage were included in the animated video. 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.

Session 2: Gustar–type verbs. For the *gustar–type verbs* structure we focused on six different verbs which follow the non-canonical argument structure and semantic mapping described in section 4.2, *gustar*‘/to like’, *encantar*/‘to love’, *interesar/*‘to be interested’,

importar/'to care', *molestar/*'to be bothered', *quedar/*'to be left'. Each of these forms was presented four times, twice in the singular form, and twice in the plural form.

Session 3: Subjunctive in noun clauses. For the subjunctive in noun clauses, five rules were targeted. Each rule was represented by four different verb instances. Twelve indicative sentences were included as fillers

Session 4: The conditional. For the conditional tense we included seven different verb which followed this construction: *estaría, bebería, visitaría, haría, vendría, podría, iría, divertiría y pasearía*. Each verb was presented either one or two times throughout the script see Supplementary Materials; Appendix B).

4.5.3. Captioning content and textual enhancement manipulations

The effect of TE on vocabulary and grammar within the captioning line was investigated through three experimental groups.

- Lesson + No Salience group (Lesson + Control) The control version of the videos did not include captions.
- Lesson + Salience on Vocabulary (Lesson +SV): The vocabulary version included captions that provided learners with textual enhancement on the target vocabulary via bold and yellow text.
- Lesson + Salience on the grammatical features group (Lesson +SG): For each of the four target structures, the grammar version included captions that provided learners with textual enhancement via bolding, color-coding or underlining on the appropriate inflectional, syntactic, and functional considerations (see Table 4.3 for a summary of these groups by structure).

Table 4.3. Summary of Captioning + Textual Enhancement manipulations per grammar topic.

	Condition		
Grammar Topic	Control	Salience on Vocabulary	Salience on Grammar
Preterite-imperfect	n/a	Target vocabulary word visually enhanced in bold and yellow	target verb is underlined; preterite target morpheme is bold and orange; imperfect target morpheme is bold and yellow
Gustar-type verbs	n/a		target verb is bold and underlined; target morpheme is coded in yellow; all other plurality markers are coded in yellow, but not in bold
Subjunctive in noun clauses	n/a		main clause verb is bold and orange, followed by an orange bold arrow; conjunction *que* 'that' in bold and white; subjunctive verb is underlined, and target subjunctive morpheme is in bold and yellow
Conditional	n/a		target verb is underlined; conditional target morpheme is bold and yellow

The No Lesson + Salience on the grammatical features group (No Lesson + SG) received the same type of TE as the Lesson + SG group. For ethical reasons, this group also received the grammar lesson but only *after* completing all of the study questions at the end of each lab session. As described previously, this group was included as a way to examine if textually enhanced captions on grammar would be facilitative in the absence of the explicit grammar lesson provided prior to the animated captioned video. For this reason, this group was analyzed separately and compared to its Lesson + SG counterpart.

Table 4.4. General overview of procedure.

Phase	Test	Time
Pre-experimental Phase	Grammar Pretest	First day of class
	Vocabulary Proficiency Test	First day of class
	Language History Questionnaire	First week of class
	Elicited Imitation Task	First Month 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.

4.6. Data collection procedure

On the first day of class of the 15-week semester, four members from the research team attended all 21 course sections and administered the Spanish vocabulary proficiency test and the

grammar pretest.[15] During the first week of class, the learners additionally filled out the web-based Language History Questionnaire through the LHQ 2.0 interface available on the Penn State Brain Language and Computation Lab website.[16] Given the large number of students, the EIT oral proficiency test was administered throughout the first month of class. All students were tested individually in a quiet room. We used Marantz Pmd620 digital recorder and Shure WH20 head-mounted microphones to conduct the recordings.

The experimental phase of the study took place over four different sessions spaced through the semester in the order presented in the class syllabus: (1) preterite/imperfect forms, (2) *gustar-type* verbs, the (3) subjunctive in noun clauses, and (4) the conditional tense. Instructors were asked not to assign readings or homework on the target material prior to the experimental sessions. During each session, the two experimenters met with the learners and instructors on their assigned class day and time, in a pre-assigned computer classroom. The experimental protocol was computerized and made available to each participant through the Canvas Learning Platform (https://www.canvaslms.com/), which allows for the creation of multimedia surveys. During each experimental session, learners in the Control, Lesson + SV and Lesson + SG group were first presented with the grammar lesson video about the target form, followed by the corresponding animated video manipulated for one of three groups: no captioning was provided (Control); target vocabulary was highlighted via TE (Lesson + SV); or grammatical features were highlighted via TE (Lesson + SG. For the No Lesson + SG group,

[15] Learners who were absent on the first day of class, or enrolled after the first week, completed the Pre-Experimental phase during a separate make-up session.

[16] The Language History Questionnaire can be accessed online through the Penn State Brain Language and Computation Lab website: http://blclab.org/language-history-questionnaire/

learners were first presented with the animated video, and saw the grammar lesson at the end of the experiment.

Following the grammar video lessons and the animated videos, participants completed four different tests, which examined their recognition and production (translation) of the target vocabulary; and their recall of form and grammar. Each session lasted approximately 50 minutes. Similar versions of the grammar and vocabulary translation tests were administered by the learners' instructors two weeks after the treatment in order to investigate retention of the targeted vocabulary and grammar structures over time (See Table 4.4 for a summary of the procedure).

4.7. Statistical Analysis

Statistical analyses were conducted using R Studio version 1.0.143 (RStudio Team, 2015). The data were analyzed by generalized linear models and multilevel generalized linear regression models utilizing the glm() and glmer() functions within the lme4 package in R (Bates, Maechler, & Bolker, 2015).

4.7.1. 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 measures were proportion of trials correct, with GROUP (Lesson + Control, Lesson + SV, Lesson + SG and No Lesson + SG). The week 1 VOCABULARY PROFICIENCY test was additionally included as a fixed variable to take into account individual differences in Spanish proficiency. This variable was mean-centered before being added to the model.

4.7.2. Grammar data

The dependent measures were proportion of trials correct, with GROUP, and STRUCTURE (preterite/imperfect, *gustar*-type verbs, subjunctive and conditional) and TIME (Pretest, Immediate Posttest and Two-week Posttest) as predictor terms, as well as random intercepts for SUBJECTS. The EIT was additionally included as a fixed variable to take into account individual differences in Spanish proficiency. This variable was mean-centered before being added to the model.

4.7.3. Missing data

Given that the learners in the current study received course credit for their participation in each of the lab sessions, they were allowed to attend a make-up session for any lab they were absent from. If participants took a make-up after being presented with the lab material by their instructor, their data for that specific lab session was treated as missing.

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 4.3.2).

5. Results

5.1. Proficiency Data

Table 4.5. Means, standard deviations (SD), and 95% confidence intervals (CI) for the vocabulary and EIT proficiency tests and the pretest recognition of target vocabulary.

Group	Mean	SD	95% CI
		Lextale-ESP Vocabulary Proficiency	
Lesson + Control	-8.338	6.019	[-8.830, -7.847]
Lesson + SV	-9.623	7.521	[-10.246, -9.000]
Lesson + SG	-8.591	6.686	[-9.111, -8.068]
No Lesson + SG	-7.253	6.412	[-7.744, -6.762]
		Elicited Imitation Task Spanish Proficiency	
Lesson + Control	66.515	19.638	[61.850, 71.178]
Lesson + SV	68.543	20.923	[63.640, 73.445]
Lesson + SG	68.819	18.102	[64.638, 73.000]
No Lesson + SG	67.908	22.430	[62.865, 72.950]

Table 4.5 presents the group means, standard deviations and confidence intervals for the Spanish Vocabulary Proficiency Test and the Elicited Imitation Task (EIT).

The vocabulary proficiency test included 46 words that were used as experimental items (23 vocabulary target words and 23 foils) in this study. These items were removed from the scoring of the proficiency test to separately assess learners' prior knowledge of these words.

5.2. Vocabulary

5.2.1. Recognition

The Vocabulary recognition data are plotted in on the left-hand panel of Figure 4.1. The pattern for the recognition data suggests an advantage of captioning over non-captioned video, with all captioning groups scoring higher than the Lesson + Control which was not presented with captions. Additionally, the data patterns suggest at an overall advantage for the Lesson + SV participants over the Lesson + Control and the two Grammar groups (Lesson + SG and No Lesson + SG) (see Supplementary Material; Appendix C; Table 4.C1 for additional details). To investigate the effects of captioning, we ran a generalized linear model which included fixed effects of VOCABULARY PROFICIENCY and our main variable of interest: GROUP. The first model, with the Lesson + Control group as the reference level, revealed significant positive group effects, for the Lesson + SV group, $\beta = 1.286$, SE = 0.075, $p < 0.001$, the Lesson + SG group, $\beta = 0.755$, SE = 0.067, $p < 0.001$, and the No Lesson + SG group, $\beta = 0.756$, SE = 0.066, $p < 0.001$. Thus, all captioned groups were more accurate in their recognition accuracy than the controls. The same model, with Vocabulary as the reference level, revealed a significant negative group effect for the Lesson + Control group $\beta = - 1.286$, SE = 0.075, $p < 0.001$; for the Lesson + SG group, $\beta = - 0.532$, SE= 0.071, $p < 0.001$; and the No Lesson + SG group, $\beta = - 0.531$, SE = 0.070, $p < 0.001$ (see Table 4.6 for details). Thus, there was an overall advantage of the Lesson + SV group in their recognition accuracy.

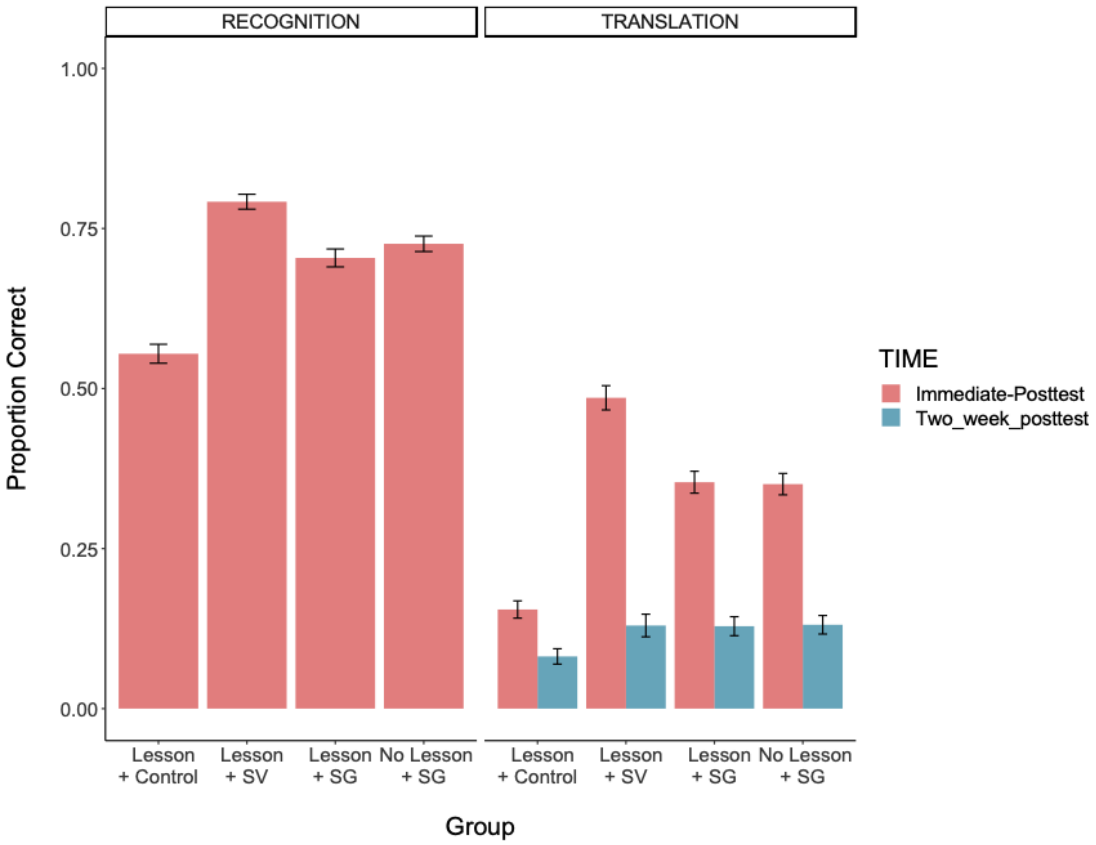

Figure 4.1.*Mean Accuracy Scores for Vocabulary (A) Recognition and (B) Translation. Error bars are 2 standard errors long.*

5.2.2. Translation

Immediate Posttest. 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 Lesson + SV group over the Control and Grammar groups (see the right-hand panel of Figure 4.1; and Supplementary Material; Appendix C; Table 4.C1). 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, for the Lesson + SV group, β = 1.528, SE = 0.099, $p < 0.001$; for Lesson + SG group, β = 1.067, SE = 0.098, $p < 0.001$; and for the No Lesson + SG group, β = 1.102, SE = 0.098, $p < 0.001$, i.e., all captioned groups were more accurate in their production accuracy. The same model, with Vocabulary as the reference

level, revealed a significant negative group effect, for Control, β = - 1.528, SE = 0.099, p < 0.001, for the Lesson + SG group, β = - 0.459, SE = 0.079, p < 0.001; and for the No Lesson + SG group, β = -.0.426, SE = 0.079, p < 0.001, confirming our initial observation of the overall advantage of the Lesson + Vocabulary group in their translation accuracy (see Table 4.7 for details).

Two-week Posttest. Similar to the immediate posttest, the pattern for the two-week posttest data suggests an advantage of captioning over non-captioned video, with all captioning groups scoring higher than the no captions Control group. However, the data pattern does not suggest an overall advantage for the Lesson + SV participants over the Grammar groups (see the right-hand panel of Figure 4.1; and Supplementary Material; Appendix C; Table 4.C2). To investigate the effects of captioning, we ran the same analysis as for the immediate posttest data. The first model, with the Lesson + Control group as the reference level, revealed significant positive group effects, for the Lesson + SV group, β = 0.464, SE = 0.207, p < 0.05, the Lesson + SG group β = 0.488, SE = 0.203, p < 0.05, and the No Lesson + SG group, β = 0.563, SE = 0.195, p < 0.01. Thus, all captioned groups were more accurate in their translation accuracy than the controls. The same model, with Vocabulary as the reference level, revealed a significant negative group effect against the Control, β = - 0.464, SE = 0.207, p < 0.05; but not against and the Lesson + SG group, β = 0.024, SE= 0.181, p = 0.895, or the No Lesson + SG group, β = 0.099, SE = 0.175, p = 0.571 (see Table 4.8 for details). Thus, at two weeks, the advantage of the Vocabulary group was only evident when compared against the no captions Control group, but not against the two captioned Grammar groups.

Table 4.6. Vocabulary Recognition result summary: coefficient estimates β, standard errors SE (β), associated Wald's z-score (= β/SE(β)) 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)	0.569	0.050	11.309	< 0.001 ***
Lesson + Vocabulary	1.286	0.075	17.194	< 0.001 ***
Lesson + Grammar	0.755	0.067	11.193	< 0.001 ***
No Lesson + Grammar	0.756	0.066	11.470	< 0.001 ***
Vocabulary Proficiency	-0.018	0.023	-0.722	0.470
Model 2 with the Vocabulary group as the reference level				
(Intercept)	1.855	0.055	33.541	< 0.001 ***
Lesson + Control	-1.286	0.075	-17.194	< 0.001 ***
Lesson + Grammar	-0.532	0.071	-7.457	< 0.001 ***
No Lesson + Grammar	-0.531	0.070	-7.566	< 0.001 ***
Vocabulary Proficiency	-0.018	0.023	-0.722	0.470

Table 4.7. Vocabulary Immediate Posttest Translation result summary: coefficient estimates β, standard errors SE (β), associated Wald's z-score (= β/SE(β)) 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)	-1.576	0.081	-19.393	< 0.001 ***
Lesson + Vocabulary	1.528	0.099	15.484	< 0.001 ***
Lesson + Grammar	1.069	0.098	10.858	< 0.001 ***
No Lesson + Grammar	1.102	0.098	11.234	< 0.001 ***
Vocabulary Proficiency	-0.067	0.029	-2.309	< 0.05 *
Model 2 with the Vocabulary group as the reference level				
(Intercept)	-0.048	0.056	33.541	< 0.001 ***
Lesson + Control	-1.528	0.099	-15.484	< 0.001 ***
Lesson + Grammar	-0.459	0.079	-5.820	< 0.001 ***
No Lesson + Grammar	-0.426	0.079	-5.410	< 0.001 ***
Vocabulary Proficiency	-0.067	0.029	-2.309	< 0.05 *

Table 4.8. Vocabulary Two-week Posttest Translation result summary: coefficient estimates β, standard errors SE (β), associated Wald's z-score (= β/SE(β)) 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)	-2.737	0.159	-17.192	< 0.001 ***
Lesson + Vocabulary	0.463	0.207	2.242	< 0.05 *
Lesson + Grammar	0.487	0.203	2.406	< 0.05 *
No Lesson + Grammar	0.563	0.195	2.884	< 0.01 **
Vocabulary Proficiency	-0.027	0.063	-0.432	0.665
Model 2 with the Vocabulary group as the reference level				
(Intercept)	-2.274	0.132	-17.231	< 0.001 ***
Lesson + Control	-0.464	0.207	-2.242	< 0.05 *
Lesson + Grammar	0.024	0.181	0.133	0.895
No Lesson + Grammar	0.099	0.175	0.567	0.571
Vocabulary Proficiency	-0.027	0.063	-0.432	0.665

5.3. Grammar: Comparing Explicit Grammar Instruction groups

5.3.1. Recall of Form

Figure 4.2 illustrates the group mean scores as well as the standard errors for all four target grammar structures (the preterite/imperfect, *gustar*-type verbs, subjunctive and the conditional). Here, the overall pattern does not suggest any clear group differences within each structure (see Supplementary Material; Appendix C; Table 4.C3). We ran a generalized linear mixed effects model, which included fixed effects of the EIT, GRAMMAR PRETEST and our main variables of interest: GROUP and STRUCTURE. The model included the Control group, and the preterite/imperfect as reference levels. We used the emmeans (Length, 2018) package to run pairwise Tukey tests examining whether there were group differences within each structure (see Table 4.9 for details). Our initial observations were confirmed by our model, which did not reveal any significant GROUP by STRUCTURE interactions (see Table 4.9 for details).

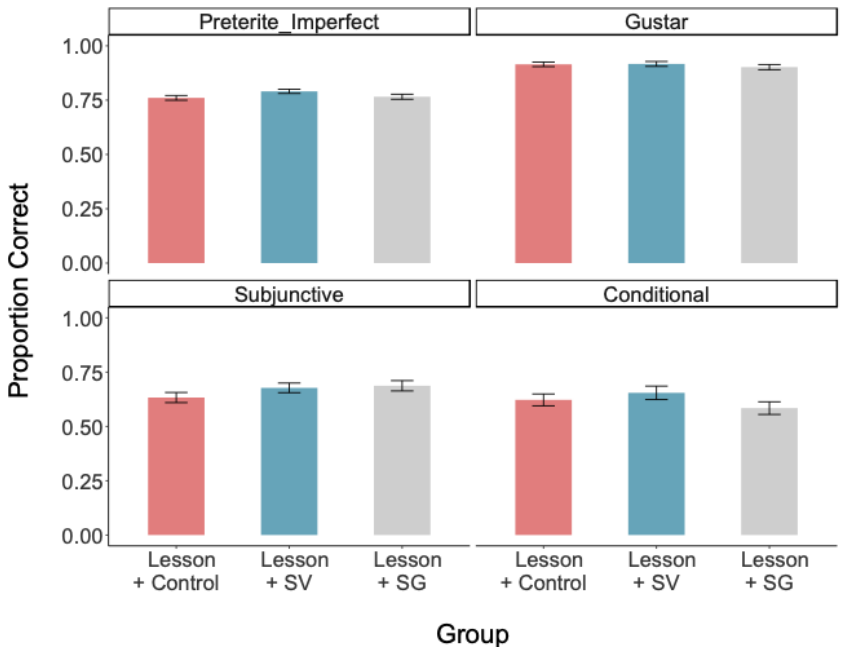

Figure 4.2. *Mean Accuracy Scores for Grammar Recall of Form by Structure and Group (Explicit Grammar Instruction groups). Error bars are 2 standard errors long.*

Table 4.9. Immediate Posttest Grammar Recall of Form emmeans contrasts summary for explicit instruction groups.

Group-Pairwise	Coef. *B*	SE	**z**	***p***
		Contrasts for the Preterite and Imperfect		
Control – Vocabulary	-0.146	0.100	-1.453	0.318
Control – Grammar	-0.460	0.099	-0.460	0.890
Vocabulary – Grammar	0.100	0.099	1.007	0.572
		Contrasts for *Gustar*-type verbs		
Control – Vocabulary	-0.047	0.142	-0.330	0.942
Control – Grammar	0.079	0.138	0.574	0.834
Vocabulary – Grammar	0.126	0.138	0.910	0.634
		Contrasts for the Subjunctive in Noun Clauses		
Control – Vocabulary	-0.181	0.124	-1.457	0.312
Control – Grammar	-0.244	0.124	-1.963	0.122
Vocabulary – Grammar	-0.063	0.125	-0.503	0.869
		Contrasts for the Conditional		
Control – Vocabulary	-0.136	0.153	-0.884	0.651
Control – Grammar	0.166	0.147	1.128	0.497
Vocabulary – Grammar	0.301	0.152	1.977	0.112

Note. Control = Lesson + Control; Vocabulary = Lesson + Vocabulary; Grammar = Lesson + Grammar. Results are given on the log odds ratio (not the response) scale. P value adjustment: Tukey method for comparing a family of 3 estimates.

5.3.2. Translation

Figure 4.3 illustrates the group mean scores as well as the standard errors by structure for the grammar Pretest, the Immediate Posttests and the Two-week Posttests. The data pattern shows similar effects by structure, whereby all groups show an increase in their immediate posttest accuracy scores when compared to their corresponding pretest scores, but no apparent differences between groups at Immediate Posttest or Two-week Posttest. In the analyses that follow, we focus on the learner groups' gains from pretest to Immediate Posttest and from Pretest to Two-week Posttest (See Supplementary Material; Appendix C; Tables 4.C4-C6).

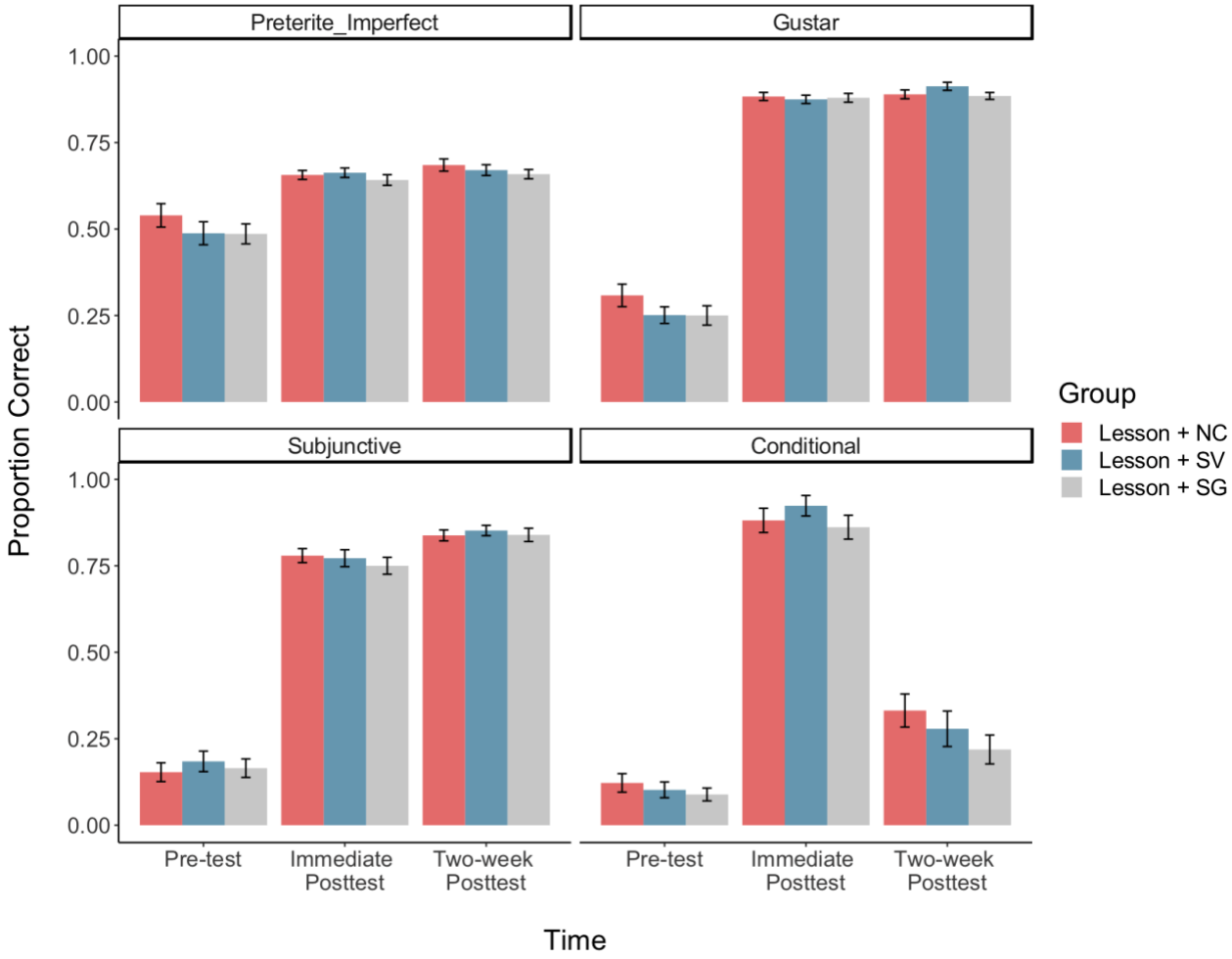

Figure 4.3. *Mean Accuracy Scores for Grammar Translation by Structure, Group and Time (Explicit Grammar Instruction groups). Error bars are 2 standard errors long.*

Table 4.10. Pretest-Immediate Posttest Grammar Translation emmeans contrasts summary for explicit instruction groups.

Time-Pairwise	Group-Pairwise	Coef. *B*	SE	**z**	***p***
		Contrasts for the Preterite and Imperfect			
Pre-test -Immediate Posttest	Control – Vocabulary	0.269	0.126	2.133	0.033
	Control – Grammar	0.228	0.126	1.813	0.069
	Vocabulary – Grammar	0.041	0.123	0.338	0.736
		Contrasts for *Gustar*-type verbs			
Pre-test -Immediate Posttest	Control – Vocabulary	0.308	0.205	1.508	0.132
	Control – Grammar	0.418	0.205	2.035	0.042
	Vocabulary – Grammar	-0.109	0.208	-0.525	0.599
		Contrasts for the Subjunctive in Noun Clauses			
Pre-test -Immediate Posttest	Control – Vocabulary	-0.246	0.194	-1.270	0.204
	Control – Grammar	-0.226	0.195	-1.160	0.246
	Vocabulary – Grammar	-0.021	0.189	-0.111	0.912
		Contrasts for the Conditional			
Pre-test -Immediate Posttest	Control – Vocabulary	0.830	0.387	2.165	0.030
	Control – Grammar	0.252	0.352	0.716	0.474
	Vocabulary – Grammar	0.586	0.390	1.502	0.133

Note. Control = Lesson + Control; Vocabulary = Lesson + Vocabulary; Grammar = Lesson + Grammar. *R*esults are given on the log odds ratio (not the response) scale. P value adjustment: Tukey method for comparing a family of 3 estimates.

5.3.2.1. Immediate Posttest

To investigate group differences, we ran a generalized linear mixed effects model which included fixed variables of EIT, and our main variables of interest: GROUP, STRUCTURE and TIME as predictor terms, with random intercepts for subjects. The model included the control group, and the preterite/imperfect structure as reference levels. We used the emmeans package to run pairwise Tukey tests comparing Pretest/Immediate Posttest gains by group within each structure (see Table 4.10 for details).

Preterite/Imperfect. The results for the preterite/imperfect revealed a significant effect in group gains between the Lesson + Control and Lesson + SV groups, $\beta = 0.270$, SE = 0.126, $p = 0.033$; a marginally non-significant effect between the Lesson + Control and Lesson + SG groups, $\beta = 0.228$, SE = 0.126, $p = 0.070$; and a non-significant effect between the Lesson + SV and Lesson + SG groups, $\beta = 0.041$, SE = 0.123, $p = 0.735$. To summarize, the Lesson + SV group led to greater translation accuracy from pretest to immediate posttest than the Lesson + Control group, whereas the Lesson + SG group showed marginal effects in the expected direction. However, there was no significant difference between the Lesson + SV and Lesson + SG groups.

Gustar-type verbs. The results for the *gustar*-type verbs revealed a non-significant effect in group gains between the Lesson + Control and Lesson + SV groups, $\beta = 0.309$, SE = 0.205, $p = 0.132$; a significant effect between the Lesson + Control and Lesson + SG groups, $\beta = 0.418$, SE = 0.205, $p = 0.041$; and a non-significant effect between the Lesson + SV and Lesson + SG groups, $\beta = -0.109$, SE = 0.208, $p = 0.599$. Thus, only the Lesson + SG group led to greater

translation accuracy from pretest to immediate posttest than the Lesson + Control group. However, there was no significant difference between the Lesson + SV and Lesson + SG groups.

Subjunctive. The results for the *gustar*-type verbs did not reveal significant effects in group gains between the Lesson + Control and Lesson + SV groups, β = - 0.259, SE = 0.194, p = 0.204; the Lesson + Control and Lesson + SG groups, β = - 0.226, SE = 0.195, $p < 0.05$; or the Lesson + SV and Lesson + SG groups, β = 0.021, SE = 0.189, p = 0.912. Thus, all groups showed similar group gains from pretest to immediate posttest.

Conditional. The results for the *gustar*-type verbs revealed a significant effect in group gains between the Lesson + Control and Lesson + SV groups, β = 0.838, SE = 0.387, $p < 0.05$; a non- significant effect between the Lesson + Control and Lesson + SG groups, β = 0.252, SE = 0.352, p = 0.474; and a non-significant effect between the Lesson + SV and Lesson + SG groups, β = 0.586, SE = 0.390, p = 0.133. Thus, only the Lesson + SV group led to greater translation accuracy from pretest to immediate posttest than the Lesson + Control group. However, there was no significant difference between the Lesson + SV and Lesson + SG groups.

5.3.2.2. Two-week Posttest

In order to investigate group differences, we ran the same model as for the immediate posttest data, followed by pairwise Tukey testing of the relevant mean scores. The results for the pairwise comparisons are shown in Table 4.11 where the only significant effect in group gains was found between the Lesson + Control and Lesson + SV groups, β = 0.630, SE = 0.210, p <0.01, for the *gustar*-type verbs structure.

Table 4.11. Pretest-Two-week Posttest Grammar Translation emmeans contrasts summary for explicit instruction groups.

Time-Pairwise	Group-Pairwise	Coef. *B*	SE	**z**	***p***
		Contrasts for the Preterite and Imperfect			
Pre-test -Two-week Posttest	Control – Vocabulary	0.149	0.128	1.164	0.245
	Control – Grammar	0.114	0.127	0.898	0.369
	Vocabulary – Grammar	0.035	0.123	0.287	0.774
		Contrasts for *Gustar*-type verbs			
Pre-test -Two-week Posttest	Control – Vocabulary	0.630	0.209	3.005	0.002
	Control – Grammar	0.289	0.205	1.413	0.158
	Vocabulary – Grammar	0.341	0.212	1.605	0.109
		Contrasts for the Subjunctive in Noun Clauses			
Pre-test -Two-week Posttest	Control – Vocabulary	-0.113	0.199	-0.569	0.569
	Control – Grammar	-0.030	0.201	-0.153	0.878
	Vocabulary – Grammar	-0.083	0.199	-0.416	0.677
		Contrasts for the Conditional			
Pre-test -Two-week Posttest	Control – Vocabulary	0.072	0.291	0.248	0.804
	Control – Grammar	-0.178	0.287	-0.622	0.534
	Vocabulary – Grammar	0.251	0.301	0.832	0.405

Note. Control = Lesson + Control; Vocabulary = Lesson + Vocabulary; Grammar = Lesson + Grammar. *R*esults are given on the log odds ratio (not the response) scale. P value adjustment: Tukey method for comparing a family of 3 estimates.

5.6. Comparing EGI + TE-captions on grammar versus No lesson + TE-captions on grammar

5.6.1. Grammar Recall of Form

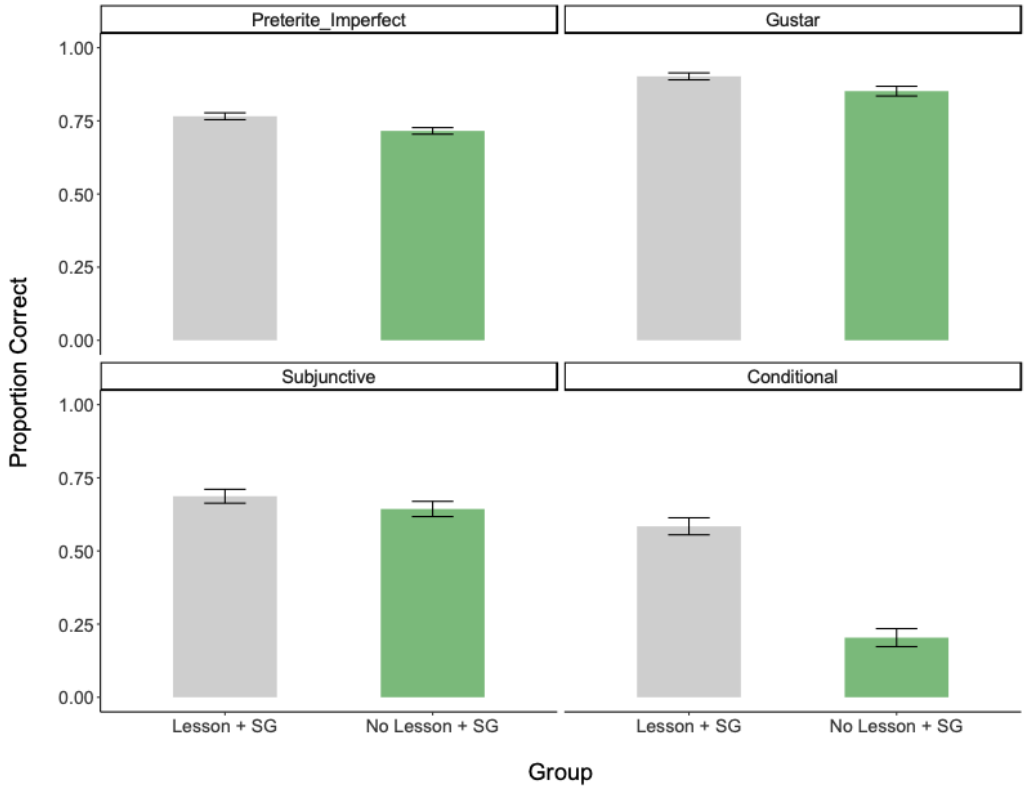

Figure 4.4. *Mean Accuracy Scores for Grammar Recall of Form by Structure and Group (Explicit Grammar Instruction versus No Lesson). Error bars are 2 standard errors long.*

Figure 4.4 illustrates the group mean scores as well as the standard errors for all four target grammar structures (the preterite/imperfect, *gustar*-type verbs, subjunctive and the conditional). The data pattern suggests an overall advantage of the explicit Lesson + SG group over the implicit No Lesson + SG group (see Supplementary Material; Appendix C; Table 4.C3 for details), especially for the Conditional tense. In order to investigate these group differences, we ran a generalized linear mixed effects model which included fixed variables of EIT, GRAMMAR

PRETEST and our main variables of interest: GROUP and STRUCTURE as predictor terms, with random intercepts for subjects. The model included the Lesson + SG group, and the preterite/imperfect structure as reference levels. We used the emmeans package to run pairwise Tukey tests examining whether there were differences between the Lesson + SG group and the No Lesson + SG groups within each structure (see Table 4.12 for details). The results revealed significant differences for the preterite/imperfect ($p < 0.001$), *gustar*-type verbs ($p < 0.001$), the subjunctive in noun clauses ($p = 0.02$) and the conditional tense ($p < 0.001$).

Table 4.12. Immediate Posttest Grammar Recall of Form emmeans contrasts summary for Explicit Grammar Instruction versus No Lesson.

Group-Pairwise	Coef. *B*	SE	**z**	***p***
		Contrasts for the Preterite and Imperfect		
Lesson + Grammar– No Lesson + Grammar	0.311	0.104	3.002	< 0.01**
		Contrasts for *Gustar*-type verbs		
Lesson + Grammar– No Lesson + Grammar	0.539	0.131	4.113	<.0001***
		Contrasts for the Subjunctive in Noun Clauses		
Lesson + Grammar– No Lesson + Grammar	0.275	0.128	2.155	<0.05*
		Contrasts for the Conditional		
Lesson + Grammar– No Lesson + Grammar	1.749	1.161	10.878	<.0001***

*Note. R*esults are given on the log odds ratio (not the response) scale. P value adjustment: Tukey method for comparing a family of 3 estimates.

5.6.2. Grammar Translation

Figure 4.5 illustrates the group mean scores as well as the standard errors by structure for the grammar Pretest, the Immediate Posttests and the Two-week Posttests. The data pattern shows similar effects by structure, whereby all groups show an increase in their accuracy when compared to their corresponding pretest scores. A closer examination of these pretest/posttest gains by group reveals varying effects by structure. From pretest to immediate posttest, the explicit Lesson + SG group appears to show a slight advantage over the No Lesson + SG group for the preterite/imperfect, the *gustar*-type verbs, and a considerable advantage for conditional tense, but not for the subjunctive in noun clauses where both groups show similar gains in their accuracy scores. These differences do not appear to hold true for the prestest/two-week posttest gains (see Supplementary Material; Tables 4.C4-C6 for details)

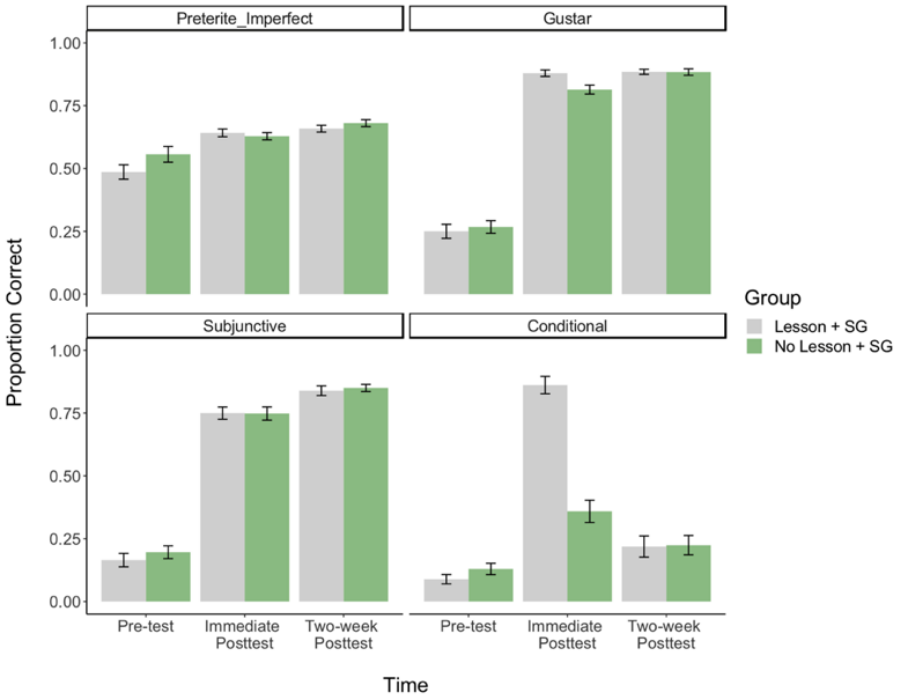

Figure 4.5. *Mean Accuracy Scores for Grammar Translation by Structure, Group and Time (Explicit Grammar Instruction versus No Lesson). Error bars are 2 standard errors long.*

Table 4.13. Pretest-Immediate Posttest grammar translation emmeans contrasts summary for Explicit Grammar Instruction versus No Lesson.

Time-Pairwise	Group-Pairwise	Coef. *B*	SE	**z**	***p***
		Contrasts for the Preterite and Imperfect			
Pre-test - Immediate	Lesson + Grammar– No Lesson + Grammar	-0.406	0.119	-3.388	< 0.01**
		Contrasts for *Gustar*-type verbs			
Pre-test - Immediate	Lesson + Grammar– No Lesson + Grammar	-0.689	0.199	-3.465	<0.01**
		Contrasts for the Subjunctive in Noun Clauses			
Pre-test - Immediate	Lesson + Grammar– No Lesson + Grammar	-0.238	0.183	-1.306	0.192
		Contrasts for the Conditional			
Pre-test - Immediate	Lesson + Grammar– No Lesson + Grammar	-3.020	0.314	-9.604	<.0001***

*Note. R*esults are given on the log odds ratio (not the response) scale. P value adjustment: Tukey method for comparing a family of 3 estimates.

Table 4.14. Pretest-Two-week Posttest grammar translation emmeans contrasts summary for Explicit Grammar Instructions versus No Lesson.

Time-Pairwise	Group-Pairwise	Coef. *B*	SE	**z**	***p***
			Contrasts for the Preterite and Imperfect		
Pre-test -Two-week	Lesson + Grammar– No Lesson + Grammar	-0.217	0.120	-1.808	0.070
			Contrasts for *Gustar*-type verbs		
Pre-test -Two-week	Lesson + Grammar– No Lesson + Grammar	-0.110	0.202	-0.551	0.582
			Contrasts for the Subjunctive in Noun Clauses		
Pre-test -Two-week	Lesson + Grammar– No Lesson + Grammar	-0.170	0.193	-0.885	0.376
			Contrasts for the Conditional		
Pre-test -Two-week	Lesson + Grammar– No Lesson + Grammar	-0.531	0.277	-1.92	0.054

*Note. R*esults are given on the log odds ratio (not the response) scale. P value adjustment: Tukey method for comparing a family of 3 estimates.

5.6.2.1. Immediate Posttest

In order to investigate group differences, we ran a generalized linear mixed effects model which included fixed variables of EIT, and our main variables of interest: GROUP, STRUCTURE and TIME as predictor terms, with random intercepts for subjects. The model included the Lesson + SG group, and the preterite/imperfect structure as reference levels. Again, we used the emmeans package to run pairwise Tukey tests comparing Pretest/Immediate Posttest gains by group within each structure. The results revealed a significant effect in group gains between the Lesson + SG and the No Lesson + SG groups for the preterite/imperfect, β = - 0.407, SE = 0.120, $p < 0.01$; *gustar-type* verbs, β = - 0.689, SE = 0.199, $p < 0.01$; and the conditional, β = -3.020, SE = 0.314, $p < 0.001$; but not for the subjunctive in noun clauses, β = - 0.238, SE = 0.182, $p = 0.192$ (see Table 4.13 for details).These results are consistent with our initial observations, whereby the Lesson + SG group showed a greater advantage for all structures, except the subjunctive.

5.6.2.2. Two-week Posttest

In order to investigate group differences, we ran the same model as for the immediate posttest data followed by pairwise Tukey testing of the relevant mean scores. We did not uncover any significant differences in pretest/two-week posttest group gains for any of the grammar structures (see Table 4.14 for details). The results thus confirm our initial observations.

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 in all three captioning groups (Lesson + Vocabulary; Lesson + Grammar; No Lesson + Grammar) were more successful than non-captioned control learners in acquiring the target vocabulary words. There was an effect of vocabulary TE on both the recognition and production scores. This is evidenced by the advantage of the Vocabulary group over both the Control and Grammar conditions – which did not include highlighting on vocabulary. The overall findings for the immediate posttest vocabulary data lend support to previous research demonstrating the role of captioning in promoting learner knowledge of L2 vocabulary (e.g., Montero-Perez, 2013) and replicate the findings for vocabulary from Study 1 (see Chapter 2 of this dissertation).

A second research aim was to investigate if any initial gains of full captions + TE vocabulary on the production of vocabulary were maintained over time. To address this question, we tested participants' ability to translate the target vocabulary words approximately two weeks after each lab session. Upon inspection of the data, there was a noticeable dip in learners' ability to produce the target vocabulary words across each experimental group. The results revealed an advantage for each captioned group (Lesson + SV; Lesson + SG; No Lesson + SG) against the Lesson + Control group, but no significant differences between the captioned groups. Our results thus partially confirm prior findings (Baltova, 1999; Neuman and Koskinen, 1992), showing that captioned media can lead to some amount of retention, however, in the context of our study, given learners' reduced ability to produce the target words, it is not clear if these effects would lead to long-term word retention.

There are several potential explanations for the lack of robust retention effects. First, the target vocabulary selected for this experiment was of low frequency – this was a characteristic of

our design in order to control for learner familiarity of the target vocabulary. Within L2 acquisition, vocabulary size, i.e., the total number of words known and that are represented in learners' long-term memory, is largely dependent upon the relative frequency to which they are encountered in the input, with higher frequency words usually making it into long-term memory earlier than low-frequency words (Nation, 2006; Webb & Nation, 2017). Additionally, in the current study, although the design provided learners with frequent and meaningful encounters to the target words within the animated videos, learners were only exposed to the target vocabulary during their designated lab sessions–they did not receive instruction on the vocabulary words nor were they encouraged to subsequently use these words throughout the semester. It is thus possible that the lack of additional opportunities to revisit the target vocabulary, in different forms and learning contexts, contributed to learners' reduced ability to accurately produce them during the delayed posttest. Future studies investigating the usage of captioned media in the classroom should thus consider the potential effect of variables such as relative frequency, instruction (e.g., Pujadas & Muñoz, submitted), and varied encounters with, and varied use of the target vocabulary on learner acquisition. As Webb and Nation assert (2017) "the greater the number of encounters [referring to vocabulary] … and the deeper the quality of the encounters, the more likely learning is to occur (p. 63).

One additional limitation that should be addressed in future research is the lack of additional assessment measures included in the delayed posttest. Our delayed-posttest only tested learners' ability to produce the target words, thus limiting our conclusions to learners' productive knowledge of the target vocabulary. As Ortega (2009), asserts, "it is typically found that learners know more words receptively than productively, particularly if they are infrequent or difficult words" (p. 88). Indeed, this observation reflects our data pattern for the recognition and

production data at Immediate Posttest. Thus, future studies should incorporate a variety of assessment measures at each time point, focused on both receptive and productive knowledge in order to better inform how captions might facilitate vocabulary development and retention (see for instance, Montero-Perez, 2014; Sydorenko, 2010).

6.2. Grammar

6.2.1. Comparing Explicit Grammar Instruction groups

Our third 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 were maintained over time in grammar production (this was our fourth research aim). Contrary to the vocabulary findings, the findings for grammar were mixed. For the recall of form task, no significant differences were found between the groups for any of the structures, whereas for the production data, captioned videos – either on vocabulary or grammar– showed an advantage over non-captioned videos for some structures but not others. Only *gustar*-type verbs revealed a significant advantage between the Lesson + SG group and the Lesson + Control group at Immediate Posttest. The results for the two-week posttest revealed a significant difference in group gains between the Lesson + SV and the Lesson + Control groups for the *gustar*-type verbs structure only.

In the sections that follow, we focus on specific effects for each structure.

Preterite and imperfect. Upon close inspection of Figure 4.3, all groups appeared to have more baseline knowledge of the preterite/imperfect than the other structures included in this study (*gustar*-type verbs, subjunctive in noun clauses and the conditional). However, when

visually comparing the pretest to immediate posttest gains for all structures, the data only show a slight increase in learners' production accuracy for the preterite/imperfect. Thus, the initial advantage of the preterite/imperfect did not seem to lead to larger learning gains. Nonetheless, the results uncovered significant differences in group gains for the Lesson + SV group and marginal effects in the expected direction for the Lesson + SG group.

Regarding the small learning gains observed for the preterite/imperfect, there are several possible explanations, namely (i) that of learners' prior knowledge of the structure, and (ii) the amount of structures being targeted during a single lab session. As mentioned in Study 1 (see Chapter 2), although little is known about the degree of prior knowledge learners require in order to benefit from TE manipulations, in their meta-analysis on TE and grammar learning, 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. Future investigations are thus needed in order to address the possible interactions between learners' prior knowledge and the nature of target forms in question. Such studies would allow us to gain greater insight regarding the appropriate degree of prior knowledge needed for successful TE interventions.

One additional explanation relates to the amount of structures being targeted in the intervention. As mentioned in our discussion of Study 2 (Chapter 3), the few studies examining the effects of TE on learner acquisition of the preterite/imperfect have yielded mixed findings (e.g., Cintrón-Valentín et al., under review (Study 1; Chapter 2); Jourdenais et al., 1995; Leeman et al., 1995; Overstreet, 1998). Overstreet specifically, discusses that the lack of a TE effect may be due to the added difficulty of learning how two forms function in contrast to one form within a specific semantic context. 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. In the case of the preterite/imperfect, at the local level, TE on these forms might actually distract learners' attention from the surrounding discourse, which offers critical information about the specific contexts in which each of 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 forms being targeted, each of which contains their own set of rules (see Appendix A, Chapter 2), and the added TE could have served as a distraction to learners. This could explain why we found more robust effects for the Lesson + SV group which received captions but did not include highlighting of the morphological forms.

Future research focusing on the effects of TE-captioned media on the preterite/imperfect would benefit from investigating if presenting these grammatical forms one at a time versus the contrast between the two during the same intervention would lead to greater learning gains. At the same time, given the importance of the surrounding discourse in understanding how each of these forms are used, it would additionally be beneficial for future work to investigate if increasing the sources of explicit information for these structures at various strategic points during the processing of the materials would lead to more robust learning outcomes. For instance, learners could be reminded– at various points throughout the presentation of the TE-captioned video–of the rules that are being presented and what the TE aims to provide.

Gustar-type verbs. For *gusta*r-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, correct 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. Once acquired, learners need only apply the same rule to each verb instance. One possible explanation for the results revealed 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. One additional explanation relates to the nature of the experimental items included in the study. Specifically, Cerezo et al. (2016) categorize *gustar*-type structures 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 their framework, Type 1 *gusta*r-type structures are considered the least complex and consist of at least three processing steps, whereas Type 4 *gusta*r-type structures are considered the most complex and consist of six or more processing steps. The former is made up of structures where the experiencer is a first-person noun or second-person singular pronoun whereas the latter includes two or more nouns as the experiencer. In the current study, the majority of the experimental sentences presented in the animated videos as well as those included in the assessment tasks fall under the Type 1 category described by Cerezo et al. (2016). It is thus possible that the specific *gustar*-type structures included in the present experiment were more easily processed by the learners leading to more robust effects of the multimodal captioned media compared to other structures. As an avenue for further investigation, future studies could well apply the type of framework described by Cerezo et al. (2016) to more thoroughly investigate the scope of TE + captions in facilitating the acquisition of diverse formsof *gusta*r-type structures.

Subjunctive in noun clauses. The results for the subjunctive did not reveal significant differences in learning gains between groups. Thus, although all groups showed a notable increase in their ability to produce the subjunctive from Pretest to Immediate Posttest, they all appeared to be performing at the same level. These results were unexpected given the findings discussed in Studies 1 (Chapter 2) and 2 (Chapter 3). Two potential explanations for these contradictory findings could be related to (i) the different grammar TE manipulations included in Studies 1 and 3 (the present experiment); as well as (ii) the different assessment measures included in Studies 2 and 3. As described in section 2.2. of the present experiment, one of the limitations of Study 1 was that it did not assess whether different types of TE have differential effects on L2 learning. Study 2 addressed this concern by assessing effects of different designs of TE video captions on learners' immediate uptake of *gustar*-type verbs, the preterite/imperfect contrast, and the subjunctive. In Study 2, the strongest effects were found for the subjunctive, where it was uncovered that providing learners with differential highlighting on both syntactic and inflectional cues (i.e., TE2) led to increased accuracy of TE2 over TE1 –which did not make such distinctions–and the non-captioned condition. Given the superior performance of TE2, we incorporated this manipulation for the subjunctive in the present experiment. However, it is possible that the effects uncovered in Study 2 were also associated to the particular assessment measure included in the study. Specifically, while Study 2 involved a measure of learners' ability to immediately reproduce the grammatical forms, the present study measured learners' ability to more explicitly apply their knowledge of the word forms through traditional recognition and production measures (see Han et al., 2008). These differences could have led to the inconsistent findings uncovered here. This will be further discussed in Chapter 5 where I offer a general discussion of all three studies.

Conditional tense. The findings for the conditional revealed a significant difference in learning gains from Pretest to Immediate Posttest between the Lesson + SV group and the Lesson + Control group. All groups, however, showed learning gains, but unlike the other structures, there was a notable drop in their performance at Two-week Posttest. As mentioned previously, conditional sentences are considered to be highly complex structures in both L1 and L2 acquisition, not only because of their morphosyntactic complexity, but also given the semantic complexity involved in learners' processing of this form (e.g., López Ornat, 1994). At the same time, we specifically targeted a low frequency usage of the conditional (see section 4.2.) whose analysis is largely dependent on the surrounding discourse. In the present study, learners were required to adequately analyze this surrounding discourse, (i) in order to understand how the structure works from the presentation of the animated video; and (ii) in order to provide the appropriate tense (i.e., preterite, imperfect, present or conditional) in the recall of form and production assessments. It is possible that similar to the preterite/imperfect structure, TE on the grammatical forms might have slightly distracted learners' attention from the critical surrounding discourse. This might explain the slight advantage of the Lesson + SV group whose TE manipulation only included highlighting of the target vocabulary and never appeared in the same sentential contexts as the target grammar. Our current experimental design, however, does not allow us to further explore this possibility. Thus, as an avenue for further investigation, it would be beneficial to directly compare enhanced captions versus unenhanced captions (i.e., simple captions not including textual enhancement manipulations) in order to assess the unique contributions of captioned media in facilitating learner acquisition of the target grammatical forms.

6.2.2. Comparing EGI + TE-captions on grammar versus No lesson + TE-captions on grammar

Our fifth research aim was to examine whether effects of full captions + TE grammar were equally facilitative in the absence of explicit instruction. RQ.5a examined the effects of full captions + TE grammar on improving learner knowledge of grammar; RQ.5b investigated if any initial gains of full captions + TE grammar are maintained over time in grammar production (this was our fourth research aim). For the recall of form task, the Lesson + SG groups and the No Lesson + SG groups were significantly different for all structures, with an advantage for the EGI group. At Immediate Posttest, the Lesson + SG showed a significant advantage for all structures except the subjunctive, however no differences were found at Two-week Posttest.

The advantage of the EGI (Lesson + SG) group over the No Lesson + SG group is not surprising and supports prior SLA research. For instance, in their meta-analysis of the effects of grammar instruction, Norris and Ortega (2000) showed that learners who received explicit types of L2 instruction outperformed learners who received implicit types. Likewise, in a more recent meta-analysis on the effects of FFI on diverse types of structures, Spada and Tomita (2010), found a larger advantage of explicit instruction in the acquisition of both complex and simple language forms (see Indrarathne & Kormos, 2017 for a more recent study). Nonetheless, despite the overall advantage for explicit instruction, Spada and Tomita (2010), report reliable small to medium effect sizes for implicit instruction, suggesting that it can be effective in certain cases. These findings, along with more recent research (e.g., Tolentino and Tokowicz, 2014) suggest that the nature of the form in question might determine the degree of instructional support required for successful acquisition.

Our data support the aforementioned conclusion. Specifically, upon close inspection of the data presented here, an interesting observation was that for all structures, except the conditional, the difference in learning gains between the Lesson + SG group and the No Lesson + SG group was only minimal. For the conditional, for which learners did not have a great amount of prior knowledge to draw upon during input processing, captions + TE alone were sufficient to produce significant learning gains from Pretest to Immediate Posttest, but to a much lesser degree than the Lesson + SG group. Thus, one additional conclusion from the current findings is that the degree to which captions + TE alone may prove facilitative in grammar development may be dependent upon learners' prior experience with the forms in question.

A limitation that should be addressed in future research, however, is the lack of additional comparison groups that did not receive explicit instruction. Specifically, we were not able to include a No Lesson + No Captions group or a No Lesson + unenhanced captions group (i.e., a group with simple captions without enhancement) due to issues of power. The inclusion of such groups would allow for more definitive conclusions regarding the effects TE-captioned media on the structures in question.

7. Conclusion

The findings of the present study demonstrate that captioning is reliably effective for vocabulary acquisition and can be helpful for the acquisition of some grammar structures. However, its effectiveness for grammar development may be mediated by the nature of the structure in question, learners' prior familiarity with the structure, and the degree of instructional support provided to the learner, that is, whether learners are provided with explicit instruction or captions + TE alone. The implications of this research will be further discussed in Chapter 5 where I offer a general discussion of all three studies presented in this dissertation.

8. References

Baltova, I. (1999). Multisensory language teaching in a multidimensional curriculum: The use of authentic bimodal video in core French. *The Canadian Modern Language Review, 56*(1), 32-48.

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*(4), 471-508.

Bardovi-Harlig, K. (2000). *Tense and aspect in Second Language Acquisition: Form, meaning, and use*. Oxford: Blackwell.

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

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

Borrás, I., & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *The Modern Language Journal, 78*(1), 61-75.

Bowden, H. W. (2016). Assessing second language oral proficiency for research. *Studies in Second Language Acquisition, 38,* 647-675.

Brysbaert, M. (2013). Lextale_FR a fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica, 53*(1), 23-37.

Canteli Dominicis, M., & Reynolds, J. J. (2014). *Repase y escriba: Curso avanzado de gramática y composición* (7th ed.). St. John's University: New York.

Cerezo , L. , Caras , A. , & Leow , R. P . (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*(2), 265-291.

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *Modern Language Journal, 80,* 183-198.

Chun, D. M., & Plass, J. L. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology, 1*(1), 1-35.

Cintrón-Valentín, M.C., García-Amaya, L, Ellis, N.C. (under review). Captioning and grammar learning in the L2 Spanish classroom. *The Language Learning Journal.*

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*(2), 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, 1284.*

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.

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.

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. New York, NY: Cambridge University Press.

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*(2), 164-94.

Ellis, N. C. (2017). Salience in usage-based SLA. In S. Gass, P. Spinner, & J. Behney (Eds.), *Salience in Second Language Acquisition* (pp. 21-39). New York: Routledge.

Ellis, R. (2012). Form-focused instruction and second language learning. *Language Teaching Research and Pedagogy* (pp. 271-306). US: Wiley-Blackwell.

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.

Gass, S. M., Spinner, P., & Behney, J. (Eds.). (2017). *Salience in second language acquisition*. New York: Routledge.

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), 1-50.

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods, 45*(3), 765–771.

Gudmestad, A. (2012). Acquiring a variable structure: An interlanguage analysis of second-language mood use in Spanish. *Language Learning, 62*(2), 373-402.

Han, Z., Park, E. S., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics, 29*(4), 597-618.

Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions. *Studies in Second Language Acquisition, 38*(3), 401-430.

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica, 35*(1), 49-66.

Jones, L., & Plass, J. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal, 86*(4), 546-561.

Jourdenais, R. M., Ota, S. Stauffer, B. Boyson, and 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: University of Hawaii Press.

Jung, U. O. (1990). The challenge of broadcast videotex to applied linguistics. *IRAL-International Review of Applied Linguistics in Language Teaching, 28*(3), 201-220.

LaBrozzi, R. M. (2016). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research, 20*(1), 75-91.

Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning, 26*, 125-134.

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*(3), 307-331.

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, B., & 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. *The Modern Language Journal, 102*(3), 557-577.

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* (Tech. Rep. No. 9, pp. 217–258). Honolulu: University of Hawai‘i, Second Language Teaching and Curriculum Center.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods, 44*(2), 325-343.

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*(3), 673-680.

Liskin-Gasparro, J. (2000). The use of tense-aspect morphology in Spanish oral narratives: Exploring the perceptions of advanced learners. *Hispania, 83*(4), 830-844.

Long, M. H. (2006). *Problems in SLA*. Mahwah, NJ: Erlbaum.

López Ornat, S. (1994*). La adquisición de la lengua española.* Madrid: Siglo XXI de España Editores.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41*(3), 720-739.

Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology*, *18*(1), 118-141.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review, 63*(1), 59-82.

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

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis.*Language Learning, 50*(3), 417-528.

Ortega, L. (2009). *Understanding Second Language Acquisition*. New York: Routledge.

Ortega, L., Iwashita, N., Rabie, S., & Norris, J. M. (1999). A multilanguage comparison of measures of syntactic complexity [Funded project]. Honolulu: University of Hawai’i, National Foreign Language Resource Center.

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.

Pujadas G., & Muñoz, C. (submitted). Incidental and intentional L2 learning through L1 and L2 subtitles.

RStudio Team (2015). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA. URL: http://www.rstudio.com/.

Sanz, C., and Morgan-Short, K. (2004). Positive evidence versus explicit rule presentation and explicit negative feedback: A computer-assisted study. *Language Learning, 54*(1), 35-78.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Seibert Hanson, A., & Carlson, M. T. (2014). Global effects: The roles of first language and proficiency in L2 processing of Spanish clitics. *Language Learning, 64*(2), 310-342.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*(2), 165-179.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*(2), 263-308.

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology, 14*(2), 50-73.

Terrell, T. (1991). The role of grammar instruction in a communicative approach. *Modern Language Journal, 75*, 52–63.

Tolentino, L. C., & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning, 64*(2), 279-309.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*(2), 183-203.

Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language teaching, 43*(1), 1-37.

Vanderplank, R. (2016). 'Effects of' and 'effects with' captions. How exactly does watching a TV programme with same-language subtitles make a difference to language learners? *Language Teaching, 49*(2), 235-250.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood, NJ: Ablex.

VanPatten, B., & Oikkenon, S. (1996). Explanation vs. structured input in processing instruction. *Studies in Second Language Acquisition, 18*(4), 495-510.

VanPatten, B., Inclezan, D., Salazar, H., & Farley, A. P. (2009). Processing instruction and dictogloss: a study on object pronouns and word order in Spanish. *Foreign Language Annals, 42*(3), 557-575.

Vázquez Rozas, V. (2006). Gustar-type verbs. In J. Clancy Clements, & J. Yoon (Eds.), *Functional approaches to Spanish syntax. Lexical semantics, discourse and transitivity* (pp. 80-114). New York: Palgrave MacMillan.

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford: Oxford University Press.

Winke, P. M., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 14*(1), 65-86.

Appendix A.

Grammar Pretest

SP277 FIRST PROFICIENCY TEST

Name _____ Section_____

For each sentence please provide the appropriate Spanish translation for all of the underlined words.

	SENTENCE	ANSWER
1	I sold my house before the prices went up.	
2	I recognized my math teacher in the grocery store.	
3	He was looking for his glasses when he saw the stain on the carpet.	
4	You found my missing gloves in the car.	
5	After a long disease, her grandmother died.	
6	Marisa and Nicole graduated from Michigan and then completed their residence at Northwestern.	
7	When Linda was living in Mexico she would always write a journal.	
8	She would always work in the lab while her boyfriend talked to his friends in the bar.	
9	My father would always run to the gym while his sister walked.	
10	Lola was preparing dinner when her husband arrived home.	
11	Manuel would always sail from Valencia to Ibiza.	
12	When I was a graduate student I would always travel to Hawaii to learn Hawaiian.	
13	Mariela got married and right afterwards she bought a cat.	
14	They knew she was studying Spanish, but they didn't know she could speak so well.	
15	My brothers came home complaining today, they are bothered by my parents' sudden divorce.	
16	Don't tell the kids they worry too much! They are very responsible, and they care about their sickly grandmother.	
17	I'm so happy to be here, don't you love that view?	

18	My friend Peter doesn't like reading as much. On the contrary, he is interested in video games.	
19	Don't leave the room so quickly kids, I think that you have one exam left to prepare for tomorrow.	
20	Lourdes, explain this to us, you only like cats. What about the other animals?	
21	He told me that it is possible that he sends it tomorrow, but it will depend on the weather.	
22	Please clean yourself up, your mother wants you to come to the living room and meet the new neighbors.	
23	You may be very strict with your daughters, but it is true that you are a very good father.	
24	It is irrefutable that we have worked harder on this group project than you have.	
25	As soon as the guests arrive to the hotel tomorrow, Maria wishes that we show them the gardens first.	
26	Maria acknowledges that she is being unfair, but she knows that her girlfriend respects her wishes.	
27	Dear respected Professors, I don't understand why, but it is obvious that you give me more work than anyone else in the office.	
28	Your father and I will see you at the County Fair unless it rains.	
29	Dear boss, I know things got off to a rough start, nonetheless, Mary and I are happy that you work with us on this.	
30	You heard what the doctor said, if you just ate, he recommends that you wait half an hour before jumping into the pool again.	
31	Dear students, for the final essay, the professor suggests that you think about a topic early in the semester.	
32	Daniel and Enrique need to make more of an effort in their schoolwork because we see that they go to all of the parties but not to his classes.	
33	Sir, you may seem content with this decision, but it's clear that you want to be the lead engineer.	
34	And Elena told her husband the following: "It angers me that your sister calls me in the middle of the night to ask for money".	
35	Although I don't like wearing sneakers, it is obvious that I walk faster with them.	
36	You are only 10 points away from qualifying for this offer, but given your history, it seems to me that we can make an exception.	
37	Dear workers, my father loves your Lamborghinis. He believes that you build them by hand here, in Sant'Agata Bolognese.	
38	The weather is going to be so bad that I don't think we are going to go to the movies.	
39	Peter doesn't care what you say, he doubts that there is anyone as honest as my father.	
40	Wait for me in the lobby, as soon as I finish this call I'll come get you.	
41	I know that the girls want more independence but it is important that they understand that both of them are still underage.	

Now you will see another group of sentences in English. Read them carefully and provide the appropriate Spanish translation for all of the underlined words. This time use only ONE word in Spanish for each response.

	SENTENCE	ANSWER
1	When I was boy I used to play with toys every day.	
2	After working all day long, they probably were hungry.	
3	After the terrible results of tonight's auction, John must be devastated.	
4	Where was George when the accident happened? I wonder if he was at home.	
5	When I was a boy, we would go to the movies.	
6	After they arrived they drank lots of water. That day they must have run more than 25 kilometers!	
7	It's 11 o'clock and María hasn't arrived yet. She must be sick.	
8	He spent too much money while in Spain. I think he must have spent more than $4000.	
9	Yesterday we arrived to the MLB and Juan wasn't there. He must have been drinking coffee!	
10	Marcos is throwing up. He must be hungover.	

Please read the following sentences, and after the = sign complete the equivalent.
There are two gaps: in the first gap, you need to decide whether the verb is SER or ESTAR; in the second gap, you need to decide whether the word is one of the following: BUENO/A, MALO/A, BIEN, or MAL.

1. Esta manzana no se puede comer. = Esta manzana _____ _____.

2. Esa persona tiene un físico atractivo. = Esa persona _____ _____.

3. María siempre ayuda a los pobres. = María _____ _____.

4. Si comes mucha fruta, vivirás más tiempo. = La fruta _____ _____.

5. Marcos siempre le roba el dinero a la gente. = Marcos _____ _____.

6. Esta sopa tiene muy buen sabor. = Esta sopa_____ _____.

Continue to the next page.

7. Me encuentro fantástico = Hoy yo _____ _____.

8. Las drogas como la heroína y la cocaína destruyen el cuerpo. = Las drogas _____ _____ para el cuerpo.

9. Los niños vomitan en el avión porque hay muchas turbulencias. Los niños _____ _____ .

10. Los robos frecuentes causan grandes pérdidas en los pequeños comercios. = Robar en los pequeños comercios _____ _____.

Appendix B.

Animated Video Script for the Conditional Session

Note: Highlighted in blue are the target vocabulary items; highlighted in yellow are the conditional target verb instances.

PART 1

[la estación de policía]

1. **Sherlock:** [RING RING] 091 sí dígame.
2. **Propietaria:** ¿Policía?
3. **Sherlock:** Sí, sí esta es la oficina de policía local. Le habla el detective García.
4. **Propietaria:** Necesito que alguien venga a mi hotel . Anoche me robaron las **alhajas** de mi madre.
5. **Sherlock:** ¿Cómo se llama el hotel?
6. **Propietaria:** Es el Hotel Renaissance
7. **Sherlock**: ¿Y dónde se encuentra?
8. **Propietaria**: Está en la Avenida Villalobos, número 1.
9. **Sherlock**: Necesito un detalle más para completar mi informe. ¿En cuánto cree que estaban valoradas las joyas de su madre?
10. **Propietaria**: Estaban valoradas en 1,000,000 dólares.
11. **Sherlock:** Eso es muchísimo dinero. Estaré allí en 5 minutos. ¡Hasta pronto!
12. **Propietaria:** Le espero. Muchas gracias.

[en el hotel]

13. **Propietaria:** Detective García, gracias por venir.
14. **Sherlock:** No hay problema. A ver, dígame, entonces ayer hubo un robo en su hotel. Por casualidad, ¿vió quién lo hizo?
15. **Propietaria:** No, pero tengo 6 **huéspedes** en el hotel. Quizás fue alguno de ellos.
16. **Sherlock**: Vayamos con calma. ¿Quiénes se quedaron en su hotel ayer por la noche?
17. **Propietaria**: Se quedaron 3 parejas: Carmen y Laura (secretaria), Pablo y Lola, y Carlos y Ana.
18. **Sherlock**: Son muchos clientes para un hotel pequeño. ¿Dónde cree que están ahora?
19. **Propietaria**: Señor, pues ahora no hay nadie en el hotel pero espero que vuelvan a lo largo del día.

20. **Sherlock**: Eso espero yo también. Pues cuando vuelvan dígales que permanezcan en la recepción del hotel porque la policía quiere hablar con ellos.
21. **Propietaria**: Claro que sí, Sr. detective.
22. **Sherlock**: Por ahora son todos posibles sospechosos y no pueden ir a ninguna parte hasta que sean investigados.
23. **Propietaria**: No se preocupe. Les diré que tienen que hablar con la policía.
24. **Sherlock***: Hablemos un poco sobre cada cliente. ¿Quién es el primer cliente que se quedó en su casa anoche?*
25. **Propietaria**: La primera es *Carmen Ochoa*, una clienta habitual. No la vi en toda la noche!
26. **Sherlock**: Vamos a ver. ¿dónde cree usted que ***estaría (1)*** Carmen anoche?
27. **Propietaria**: Pues no lo sé, pero vamos a pensar. Yo sé que a Carmen le gusta tomarse un buen ***lingotazo*** antes de irse a dormir.
28. **Sherlock**: ¿Y se lo ***bebería (1)*** aquí?
29. **Propietaria**: No. Ella ***iría (1)*** a Eclipse, el **antro** que frecuenta de costumbre.

[Imagen de Carmen bebiendo sola en la barra del bar]

30. **Sherlock**: Supongo que Carmen ***bebería(2)*** muchas copas de vino anoche y se **divertiría(1)** en Eclipse.
31. **Propietaria**: Pero conozco bien a Carmen. Ella es una arquitecta muy rica y un 1,000,000 de dólares no es nada para ella.
32. **Sherlock**: Nunca se sabe, no solo se roba por dinero en esta vida ¿Fue Carmen a Eclipse con alguien?
33. **Propietaria**: Supongo que ***iría (2)*** con su acompañante, Laura.
34. **Sherlock**: ¡Eso es muy importante! Entonces Carmen no estaba sola.
35. **Propietaria**: Bueno, no. Carmen llegó con una acompañante y yo las vi salir juntas.

[Imagen de Carmen y Laura bebiendo en el restaurante]

36. **Sherlock:** Veamos, ¿y quién más se hospedó anoche en el hotel?
37. **Propietaria**: Pues una pareja, Pablo y Lola.
38. **Sherlock:** Veamos, entonces, ¿dónde cree usted que **estarían** (2) Pablo y Lola cuando ocurrió el robo, ayer por la noche?
39. **Propietaria:** Pues Pablo y Lola preguntaron por una galería de arte que tiene una exposición de Botero. La galería está abierta hasta tarde y la **visitarían(1)** .
40. **Sherlock**: Con el dinero del robo **podrían(1)** comprarse incluso un cuadro de Botero. ¿Qué más? ¿Quién más se quedó en el hotel anoche?
41. **Propietaria**: Otra pareja, Carlos y Ana.
42. **Sherlock**: ¿Y que cree que **harían(1)** ellos?

43. **Propietaria**: Ellos quieren casarse aquí, seguramente ***visitarían(2)*** **el salón de celebraciones** que tenemos en la propiedad.
44. **Sherlock**: ¿También tiene un salón de celebraciones aquí?
45. **Propietaria**: Sí, pero no está en este edificio.
46. **Sherlock**: ¿Y a cuánto tiempo está?
47. **Propietaria**: Está a 40 minutos a pie y el camino es precioso.
48. **Sherlock**: Pues **pasearían(1)** un buen rato. Con el dinero de las **alhajas** se pueden pagar los gastos de una boda.
49. **Propietaria**: No lo sé. Yo quería venderlas para remodelar mi hotel y hacerlo más grande. ¡Necesito su ayuda Sherlock! **[lágrimas]**
50. **Sherlock**: Bueno mujer, no se preocupe. Yo le ayudaré a encontrar al ladrón. Ahora hay que investigar lo que hicieron sus **huéspedes** ayer por la noche.
51. **Propietaria**: **[asiente con la cabeza]** ¡Mucha suerte Sherlock!

PART 2

[Sherlock visita en este orden el salón de celebraciones y habla con la encargada y comprueba que Carlos y Ana sí estuvieron allí ayer]

52. **Sherlock:** Buenos días, soy el detective García de la policía local. Estoy haciendo una investigación.
53. **Encargada:** Sí señor ningún problema.
54. Sherlock: Quería saber si una pareja vino aquí ayer por la noche.
55. Encargada: Pues, ayer vino mucha gente durante el día pero no por la noche. ¿Cómo se llamaban?
56. **Sherlock:** ¿Le dice algo el nombre de Carlos y Ana?
57. **Encargada:** Pues la verdad es que no, Además ¿para qué **vendrían(1)** aquí de noche? El salón cerró ayer a las 6:00 de la tarde.
58. **Sherlock:** Creo que buscaban un lugar para celebrar su boda. Quizás lo querían ver iluminado. A la gente le gusta hacer ese tipo de cosas!
59. **Encargada**: Ahora que lo dice, cuando iba para mi casa vi a una parejita tomándose un piscolabis en el jardín. Tenían puestas velas y todo.
60. **Sherlock:** Imagino que se **divertirían(2)** mucho en el jardín. ¿Son estos? **[enseña la foto en su móvil].**
61. **Encargada:** Sí sí esa es la pareja.
62. **Sherlock:** Entonces Carlos y Ana sí estuvieron aquí. Mmm, muchas gracias y hasta la próxima.

[Después se va a al museo habla con la chica de la limpieza y determina que Pablo y Ana si estuvieron participando en el tour nocturno de la galería]

63. **Sherlock:** Buenos días, soy el detective García de la policía local.
64. **Manager**: Hola, buenos días.
65. **Sherlock:** Una pregunta, ¿vio anoche a esta pareja **[saca el teléfono con la foto de Pablo y Ana]**?
66. **Manager:** Sí, ellos estuvieron aquí, toda la noche.
67. **Sherlock:** ¿Está seguro?
68. **Manager**: Sí, completamente, después de la cena, yo mismo hice de guía por las habitaciones de la galería. Ahora tenemos una exposición de Botero y anoche había un buen número de visitantes.
69. **Sherlock:** Entonces Pablo y Ana sí estuvieron aquí. Bueno, muchas gracias y hasta la próxima. Ahora voy a visitar el **antro** del pueblo.
70. **Manager:** Ohhh, pues que se divierta, detective García!

[Finalmente, Sherlock va al antro Eclipse y comprueba que Carmen no estuvo acompañada en el bar]

71. **Sherlock:** Buenas, soy el Detective García de la policía local. ¿Reconoce a esta mujer?
72. **Camarera:** Claro que sí, sigue allí sentada.
73. **Sherlock:** ¿Cómo?
74. **Camarera:** Sí, mire a su derecha. Lleva aquí toda la noche bebiendo.
75. **Sherlock: [camina hacia Carmen]** ¿Es usted Carmen Ochoa?
76. **Carmen:** Sí, soy yo. ¿Qué quiere?
77. **Sherlock:** Soy el detective García de la policía local.
78. **Carmen: [lo interrumpe**]. ¿Y por qué **vendría(2)** aquí un detective? Nadie está haciendo nada malo.
79. **Sherlock:** Precisamente ha habido un robo en el Hotel Renaissance
80. **Carmen:** ¿Sí? ¡Qué extraño!
81. **Sherlock:** Pues sí, alguien robó joyas anoche por valor de 1,000,000 de dólares.
82. **Carmen**: Eso es muchísimo dinero para un hotel tan pequeño
83. **Sherlock**: Se ve que eran joyas muy valiosas. Pero cuénteme, ¿cuánto tiempo lleva aquí?
84. **Carmen:** No lo sé, un buen rato supongo.
85. **Sherlock:** Creo que ha bebido suficiente.
86. **Carmen:** ¿Yo? Si sólo llevo un par de **lingotazos.**
87. **Sherlock:** Sí, claro, los últimos.
88. **Carmen:** Déjeme en paz! Me gusta beber. Esto es un país libre, ¿verdad? Me ayuda con la creatividad, soy una gran arquitecta. ¿Acaso no me conoce?
89. **Sherlock:** Eso no es importante ahora ¿Vino sola aquí anoche?
90. **Carmen:** No! Vine con mi secretaria, Laura, que por cierto**,** ¿qué **haría(2)** esa chica toda la noche? No la veo desde las 2:00 de la mañana?!?!?
91. **Sherlock**: Pero ya han pasado 8 horas!
92. **Carmen**: Pues desde entonces no la he vuelto a ver.

93. **Sherlock:** Aha, ya tengo a mi ladróna! **[detective music]**
94. **Carmen:** Ladrona ¿pero de qué habla usted señor?
95. **Sherlock:** Carmen, usted no está en condiciones de entender esta situación, pero de todos los huéspedes del hotel la única que no puede explicar su paradero es Laura.
96. **Carmen:** A ella siempre le ha gustado el dinero. Creo que por eso trabaja conmigo.
97. **Sherlock:** Pues ahora mismo voy a poner una orden de búsqueda y captura para Laura, su secretaría.
98. **Carmen:** Vaya hombre, pues ahora me tengo que buscar una secretaria nueva.

Appendix C.

Additional Data Tables

Table 4.C1. Group Proportion Correct Scores for Vocabulary Recognition and Translation at Immediate Posttest.

Group	Mean	SD	95% CI
	Group Accuracy scores for Recognition		
Lesson + Control	0.569	0.119	[0.541, 0.596]
Lesson + SV	0.859	0.095	[0.837, 0.881]
Lesson + SG	0.758	0151	[0.725, 0.791]
No Lesson + SG	0.788	0.106	[0.765, 0.811]
	Group Accuracy scores for Translation		
Lesson + Control	0.182	0.141	[0.149, 0.214]
Lesson + SV	0.488	0.199	[0.441, 0.534]
Lesson + SG	0.379	0.197	[0.336, 0.423]
No Lesson + SG	0.378	0.204	[0.334, 0.422]

Note. CI = confidence interval

Table 4.C2. Group Proportion Correct Scores for Vocabulary Recognition and Translation at Two-week Posttest.

Group	Mean	SD	95% CI
Lesson + Control	0.065	0.094	[0.043, 0.087]
Lesson + SV	0.085	0.115	[0.058, 0.112]
Lesson + SG	0.092	0.101	[0.070, 0.114]
No Lesson + SG	0.094	0.105	[0.071, 0.117]

Table 4.C3. Group Proportion Correct Scores for Grammar Recall of Form by Structure.

Group	Mean	SD	95% CI
Accuracy scores for the preterite and imperfect contrast			
Lesson + Control	0.760	0.094	[0.738, 0.783]
Lesson + SV	0.790	0.080	[0.772, 0.809]
Lesson + SG	0.766	0.103	[0.743, 0.788]
No Lesson + SG	0.716	0.101	[0.694, 0.738]
Accuracy scores for *gustar*-type verbs			
Lesson + Control	0.915	0.089	[0.894, 0.935]
Lesson + SV	0.917	0.094	[0.895, 0.939]
Lesson + SG	0.902	0.103	[0.879, 0.925]
No Lesson + SG	0.851	0.147	[0.819, 0.883]
Accuracy scores for the subjunctive in noun clauses			
Lesson + Control	0.633	0.192	[0.589, 0.677]
Lesson + SV	0.677	0.182	[0.635, 0.720]
Lesson + SG	0.687	0.201	[0.643, 0.731]
No Lesson + SG	0.643	0.235	[0.592, 0.694]
Accuracy scores for the conditional tense			
Lesson + Control	0.622	0.218	[0.572, 0.673]
Lesson + SV	0.655	0.231	[0.600, 0.701]
Lesson + SG	0.584	0.237	[0.532, 0.636]
No Lesson + SG	0.204	0.263	[0.147, 0.261]

Note. CI = confidence interval

.

Table 4.C4. Group Proportion Correct Scores for Pretest Grammar Translation by Structure.

Group	Mean	SD	95% CI
Accuracy scores for the preterite and imperfect contrast			
Lesson + Control	0.539	0.275	[0.476, 0.603]
Lesson + SV	0.488	0.276	[0.423, 0.552]
Lesson + SG	0.486	0.253	[0.430, 0.541]
No Lesson + SG	0.556	0.281	[0.496, 0.617]
Accuracy scores for *gustar*-type verbs			
Lesson + Control	0.308	0.274	[0.245, 0.372]
Lesson + SV	0.251	0.198	[0.205, 0.298]
Lesson + SG	0.250	0.244	[0.196, 0.304]
No Lesson + SG	0.267	0.222	[0.219, 0.315]
Accuracy scores for the subjunctive in noun clauses			
Lesson + Control	0.153	0.222	[0.102, 0.205]
Lesson + SV	0.184	0.243	[0.128, 0.241]
Lesson + SG	0.165	0.224	[0.116, 0.214]
No Lesson + SG	0.196	0.229	[0.147, 0.246]
Accuracy scores for the conditional tense			
Lesson + Control	0.122	0.208	[0.074, 0.170]
Lesson + SV	0.102	0.173	[0.061, 0.143]
Lesson + SG	0.089	0.147	[0.056, 0.121]
No Lesson + SG	0.129	0.183	[0.090, 0.169]

Note. CI = confidence interval

.

Table 4.C5. Group Proportion Correct Scores for Immediate Grammar Translation by Structure.

Group	Mean	SD	95% CI
Accuracy scores for the preterite and imperfect contrast			
Lesson + Control	0.656	0.105	[0.632, 0.680]
Lesson + SV	0.662	0.114	[0.636, 0.689]
Lesson + SG	0.642	0.136	[0.612, 0.672]
No Lesson + SG	0.629	0.129	[0.601, 0.657]
Accuracy scores for *gustar*-type verbs			
Lesson + Control	0.883	0.097	[0.861, 0.906]
Lesson + SV	0.875	0.099	[0.852, 0.898]
Lesson + SG	0.879	0.112	[0.855, 0.904]
No Lesson + SG	0.814	0.157	[0.780, 0.848]
Accuracy scores for the subjunctive in noun clauses			
Lesson + Control	0.779	0.166	[0.741, 0.818]
Lesson + SV	0.772	0.202	[0.725, 0.819]
Lesson + SG	0.750	0.207	[0.704, 0.796]
No Lesson + SG	0.779	0.166	[0.697, 0.799]
Accuracy scores for the conditional tense			
Lesson + Control	0.881	0.279	[0.817, 0.946]
Lesson + SV	0.924	0.219	[0.872, 0.975]
Lesson + SG	0.862	0.278	[0.800, 0.922]
No Lesson + SG	0.359	0.379	[0.277, 0.440]

Note. CI = confidence interval

.

Table 4.C6. Group Proportion Correct Scores for Two-week Grammar Translation by Structure.

Group	Mean	SD	95% CI
Accuracy scores for the preterite and imperfect contrast			
Lesson + Control	0.685	0.145	[0.651, 0.719]
Lesson + SV	0.670	0.129	[0.640, 0.701]
Lesson + SG	0.659	0.117	[0.633, 0.685]
No Lesson + SG	0.681	0.124	[0.654, 0.707]
Accuracy scores for *gustar*-type verbs			
Lesson + Control	0.889	0.102	[0.865, 0.914]
Lesson + SV	0.913	0.095	[0.891, 0.935]
Lesson + SG	0.885	0.084	[0.866, 0.903]
No Lesson + SG	0.884	0.111	[0.859, 0.908]
Accuracy scores for the subjunctive in noun clauses			
Lesson + Control	0.838	0.126	[0.809, 0.867]
Lesson + SV	0.852	0.118	[0.824, 0.879]
Lesson + SG	0.839	0.154	[0.805, 0.873]
No Lesson + SG	0.850	0.121	[0.824, 0.876]
Accuracy scores for the conditional tense			
Lesson + Control	0.332	0.377	[0.244, 0.419]
Lesson + SV	0.279	0.379	[0.189, 0.367]
Lesson + SG	0.219	0.342	[0.1430.294]
No Lesson + SG	0.224	0.319	[0.155, 0.293]

Note. CI = confidence interval.

.

CHAPTER 5. Conclusion

The goal of my dissertation was to examine how Form Focused Instruction (FFI) techniques in combination with captioned video might aid in facilitating grammar development. Across three studies, we found that grammar development can be facilitated by such multimodal techniques, however, the effectiveness of these techniques may be mediated by the specific assessment measures included in the studies – that is whether they measure immediate attention or require more elaborate processing; the specific grammatical forms in question; the degree of familiarity to the target structure the learner may bring to bear in the learning process; and the amount of instructional support provided to the learners. In the following sections, I will summarize the key findings of each of the studies included in this dissertation. I will additionally discuss specific implications for research and pedagogy.

1. Summary of Key Findings

1.1. Studies of long-term learning (Studies 1 and 3)

Study 1 (Chapter 2) of this dissertation provided a first investigation on the effects of FFI and captioned media in the L2 Spanish classroom. The results replicated prior findings in the captioning and vocabulary learning literature, revealing clear effects of captioning on both recognition and production assessments. The findings for vocabulary were also replicated in Study 3 (Chapter 4), where we additionally investigated long-term effects of productive

.

vocabulary knowledge, finding some evidence indicating that TE-captioning can lead to some degree of retention. Regarding our findings for grammar development through FFI + captioned video, the results for Study 1 showed that the learning of certain structures – namely that of *gustar*-type verbs and the subjunctive – can be facilitated by such techniques. However, the lack of a grammar pretest on learners' baseline knowledge of these forms makes it difficult to tease apart any possible confound regarding the gains acquired through the treatment from pre-existing knowledge. Additionally, the absence of a non-instructed group (a group which did not receive an explicit grammar lesson prior to the animated video), makes it difficult to tease apart whether the use of captioning was the single contributing factor to any positive effects in the learning assessments.

Study 3, specifically addressed the limitations of Study 1, revealing effects of captioned videos –either on vocabulary or grammar– on the learning of some structures but not others. Specifically, in comparing the groups which received an initial explicit grammar lesson prior to the presentation of the animated videos, (i) the findings for the *gustar*-type verbs structure showed that the Lesson + SG group led to greater translation accuracy from pretest to Immediate posttest than the Lesson + Control group; (ii) for both the preterite/imperfect and the conditional tense, the findings revealed a significant advantage for the Lesson + Salience on Vocabulary (SV) group against the Lesson + Control group; whereas, (iii) for the subjunctive in noun clauses, no significant differences were found between the experimental groups. The comparisons for the Lesson + SG group against the No Lesson + SG group –which did not include an initial explicit grammar lesson– showed a significant advantage for the Lesson + SG group for all structures except the subjunctive, where both groups showed similar gains in their accuracy scores. Upon close inspection of the data, an interesting observation was that the degree

.

to which captions + TE alone may prove facilitative in grammar development may be dependent upon learners' prior experience with the forms in question. For all structures, except the conditional, the difference in learning gains between the Lesson + SG group and the No Lesson + SG group was minimal. For the conditional, for which learners did not have a great amount of prior knowledge to draw upon during input processing, captions + TE alone were sufficient to produce significant learning gains from pretest to immediate posttest, but to a much lesser degree than the Lesson + SG group.

Across these two studies, we have shown that captioning is reliably effective for vocabulary acquisition and can be helpful for the acquisition of some grammar structures, although its effectiveness may be mediated by the nature of the structure in question, learners' prior familiarity with the structure, and the degree of instructional support provided to the learner, that is, whether learners are provided with explicit instruction or captions + TE alone.

1.2. Study on learner uptake (Study 2)

Study 2 (Chapter 3) addressed one additional limitation from Study 1 and the TE literature in general, namely that the design of Study 1 did not consider the relative influence of different types of TE on grammar learning. Study 2 thus assessed the effects of different designs of TE video captions on learners' immediate uptake of three grammatical constructions in L2 Spanish (*gustar*-type verbs, the preterite/imperfect, and the subjunctive). The results revealed that captions incorporating some type of TE (on the complete verb form, or on the critical morphological and grammatical cues and their relations), led to increased accuracy in learners' immediate reproductions of the target grammatical forms relative to the non-captioned control conditions. For the subjunctive specifically, TE highlighting the target grammatical morphemes

.

and their syntactic dependencies led to more accurate reproductions, whereas for the *gustar*-type verbs and the preterite/imperfect no differences were found between the two TE conditions.

2. Implications for research on TE and captioned media

Until recently, captioning research had primarily focused on its capacity to facilitate vocabulary learning and comprehension. The studies presented in this dissertation provide a first step in understanding how captioning in combination with FFI techniques might aid in grammar development. The findings of these studies also serve to illustrate the extra difficulties involved in the learning of grammar and highlight potential avenues for future research. In the sections that follow, I list a number of key variables that should be considered in future designs investigating the effects of captioned media and FFI. Within each section I discuss how the findings, and/or the shortcomings of the studies included in this dissertation can serve as a starting point for future research.

2.1. TE, noticing and learning

It is important to note that, across the three studies in this dissertation, the assessment measures used to investigate the effectiveness of TE-captioned media on grammar development varied, with Study 2 measuring learner uptake of the grammatical forms, and Studies 1 and 3 measuring learner recognition and production of the grammatical forms. While Study 2 involved a measure of learners' ability to reproduce the grammatical forms immediately following their exposure to experimental variations of textually enhanced captions, Study 3 measured learners' ability to more explicitly apply their knowledge of these forms following the presentation of the

.

animated videos, the latter being the more traditional acquisition measures used in the literature (see Han et al., 2008). This measure difference could help explain why the results of the three studies included here provide somewhat contradictory findings for certain structures when comparing the individual structure effects of Studies 1 and 3 to that of Study 2. Specifically, for the subjunctive and the preterite/imperfect, the results of the uptake study (Study 2) showed an advantage for TE-captioned media against non-captioned media in improving learner knowledge of these forms. However, this difference was not evidenced in Study 1 for the preterite/imperfect or in Study 3 for both the subjunctive or the preterite/imperfect.

Regarding learner acquisition of grammatical forms through TE, Han et al. (2008) assert:

> "What is crucial is whether they [L2 learners] can act upon the noticed features, and this would depend on whether or not they are able to perceive, store, and apply the salient information. Accordingly, it is hoped that input enhancement will spark a chain of cognitive processes initiated by noticing. Whether these processes can be, and how much time is needed for them to be, set in motion and completed has yet to be empirically ascertained." (p. 602)

Thus, although Study 2 shows TE may aid in re-focusing learner attention to notice non-salient forms in the input, it does not always follow that this noticing of forms will necessarily lead to their immediate acquisition, or more specifically, to learners' ability to apply their knowledge of these forms in traditional recognition and production measures. It is difficult, however, to draw further conclusions regarding the potential relationship between uptake and noticing, and the differing results of Study 2 versus those of Studies 1 and 3, given that none of the studies included in this dissertation explicitly investigated the potential links between

.

textually enhanced captions, attention, and L2 grammatical development within one experimental design.

As Winke (2013) asserts, understanding the underlying perceptual and cognitive processes of attention to language is fundamental in understanding language learning itself. This assertion is consonant with the *Noticing Hypothesis* (see section 2, Chapter 1) which holds that conscious attention to linguistic forms (e.g., sounds, words, grammar) in the input is an important ***precondition*** to learning. The majority of unimodal studies assessing TE and learning, however, have relied on acquisition measures only, ignoring how TE may serve to guide learner attention to linguistic forms and promote subsequent learning. Only a few studies (e.g., Cintrón-Valentín & Ellis, 2015; Indrahane & Kormos, 2017; Simard & Foucambert, 2013; Winke, 2013) investigate learners' noticing while exposed to the experimental input, and whether more noticing leads to greater learning gains. These studies, have, for instance, included eye-tracking to measure learners' visual attention to form, capturing noticing as it unfolds (for a methodological overview, see for instance, Roberts & Siyanova-Chanturia, 2013), with some finding strong links between attention and subsequent acquisition (e.g., Cintrón-Valentín & Ellis, 2015; Indrahane & Kormos, 2017), whereas others have not (Simard & Foucambert, 2013; Winke, 2013).

The two studies that found strong links between degree of noticing and acquisition (Cintrón-Valentín & Ellis, 2015; Indrahane & Kormos, 2017), included additional instructional tools that may have facilitated in improving learner knowledge of the forms in question. In Cintrón-Valentín and Ellis (2015), for instance, the TE condition included both visual salience and corrective feedback, whereas in Indrahane and Kormos (2017), attention and learning was most associated in the enhancement group which included explicit explanation. These groups

.

might thus have benefitted from the provision of compound enhancement, that is, "TE in combination with attention-getting strategies such as corrective feedback" (Han et al., 2008, p. 609) which tends to be more effective in encouraging noticing and subsequent processing than simple enhancement. Neither Simard and Foucambert (2013) nor Winke (2013) provided such additional instruction: Perhaps learners need to be shown why TE is being provided and how it might help.

Future research on the effects of FFI and multimodal media (e.g., captioned video) on grammar learning should therefore consider the inclusion of research tools designed to measure learners' immediate noticing of perceptually enhanced input in addition to more traditional acquisition measures. The inclusion of research tools such as eye-tracking (see for instance Lee & Révész, 2018; Montero Perez, Peters & Desmet, 2015; Muñoz, 2017) would allow for a more complete understanding of the potential interaction of salience, learner attention and TE-captioned video in L2 grammar development.

2.2. The nature of the grammatical forms

Research examining the effects of TE on L2 grammar acquisition has yielded mixed findings with some studies suggesting that its efficacy may be modulated by the linguistic form in question (e.g., Comeaux & McDonald, 2017; Leow et al., 2013) and by the type of TE that is used (LaBrozzi, 2016). Critically, these studies have compared the effects of different TE manipulations (e.g., upper-case versus lower case), but do not focus on tailoring the TE to a target morpheme (e.g., past-tense suffix) in comparison to tailoring the TE to a full lexical entry (e.g., the full verb form that contains the target morpheme). Study 2 addressed this gap in the literature by assessing which designs of TE are optimal for focusing learner attention on different linguistic constructions. Study 2 suggests that the optimal design of the TE manipulation – be it

.

focused on a full lexical entry or the target morpheme and/or additional sentential cues – should be carefully tailored to the target structure in question. Future studies investigating the effects of TE in both unimodal and multimodal contexts should employ similar designs in order to better understand the type and amount of TE that is necessary to induce noticing and promote subsequent learner knowledge of the targeted forms in question. These studies should additionally consider the potential effect of TE in adjacent versus non-adjacent dependencies (see section 2.1 in Chapter 3) in more hypothesis-driven designs.

2.3. Degree of prior knowledge

Although little is known about the degree of prior knowledge that learners require in order to benefit from TE manipulations, in their meta-analysis on TE and grammar learning, both Lee and Huang (2008) and Han et al. (2008) suggest that it may serve to moderate the effectiveness of TE. Specifically, in their meta-analysis on the effects of TE on learning, 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. Additionally, in their review on the effectiveness of TE, Han et al. (2008) suggest that the amount of learners' prior knowledge of the target forms in question may serve to determine the ***amount of instructional support*** required by the learner, that is, whether simple enhancement (i.e., TE alone) or compound enhancement (TE + in combination with any other attention-getting or instructional strategy) would be most beneficial.

In our Studies 1 and 3, we believe learners' prior knowledge or experience with the *ser* and *estar* contrast and the conditional tense, respectively, may have modulated the effects uncovered for each of these structures. Given that we did not include a measure of learners' baseline knowledge for the target grammatical forms for Study 1, any conclusion regarding

.

learners' prior experience with the *ser* and *estar* contrast is merely speculative. However, given the typical L2 developmental trajectory of the *ser* and *estar* contrast (see for instance VanPatten, 1987) and the general proficiency level of our learners, it is possible that the amount of learners' prior knowledge for this structure coupled with the initial grammar lesson, could have led to the near-ceiling effects uncovered for all three experimental groups.

With regards to the conditional structure discussed in Study 3, we targeted a very specific low frequency usage of this construction which deviated from the usage either included in the course textbook or in regular class discussion. When observing the data pattern for the conditional, Figure 4.3 in Chapter 4, which illustrates all groups that received an initial grammar lesson, showed: (i) all groups had minimal baseline knowledge of the conditional (below .15); and (ii) all groups showed notable learning gains from pretest to posttest; (iii) and these learning gains were all similar in magnitude. When observing the data pattern for this structure in Figure 4.4 in Chapter 4, which illustrates the comparison between the Lesson + SG group and the No Lesson + SG group, the data showed: (i) that both groups displayed learning gains from pretest to posttest; however (ii) the learning gains for the Lesson + SG group were greater in magnitude. This observation is contrary to the other structures, where the difference in learning gains between the Lesson + SG group and the No Lesson + SG group was minimal.

These findings suggest that TE-captioned media alone might be more helpful for certain forms, such as *gustar*-type verbs, than others, whereas for structures such as the conditional, learners might require more guided techniques in order to promote more successful acquisition of these forms, for example, by providing explicit instruction prior to the presentation of TE-captioned videos. Future studies on FFI and captioned media should thus consider, not only the nature of target forms in question, but also the amount of prior knowledge learners may bring to

.

bear. Such considerations will influence the amount of instructional support that might be required for learners to benefit from TE-captioned media.

There are several other research priorities that should be considered in future investigations. Future research should directly compare enhanced captions versus unenhanced captions (i.e., simple captions not including textual enhancement manipulations) in order to assess the unique contributions of captioned media in facilitating learner acquisition of the target grammatical forms. None of the experiments presented in this dissertation included a simple captioned condition, thus we were not able to tease apart any confounding effects of the written modality of captioning itself, from the incorporation of TE in addition to captioning. Furthermore, future designs should additionally take into account the relationship between input modality and test modality on learner outcomes. For example, Sydorenko (2010), in a study on L2 vocabulary learning, presented English-speaking L2 learners of Russian with various experimental conditions that differed in the degree of aural and visual support included in the videos (aural support only, written support only or both aural and written support), and subsequently tested their aural and written recognition of the targeted vocabulary forms. Sydorenko (2010) found differential effects on learning, according to the modality of input presentation: (i) learners who received some combination of written support scored higher on written than on aural recognition of word forms; whereas (ii) learners who received aural support only scored higher on aural than on written recognition of word forms. In the present dissertation, we only included written outcome measures, and thus cannot infer what effects, if any, the TE-captioned media had on learners' aural competence. Finally, future studies including a battery of measures ranging in their implicitness/explicitness (Norris & Ortega, 2000), would

.

allow us to develop a more complete understanding of grammar development through the usage of captioned media.

3. Implications for pedagogy

The findings of the present dissertation offer specific implications for pedagogical practices in the L2 classroom: Captions + TE can be a useful tool for both learners and L2 instructors, but the successful integration of such techniques in the L2 Spanish classroom warrant: (1) explanation of what the TE aims to provide; (2) more guided instruction for certain structures; and (3) a more prominent role of the learner in order to make L2 development a more active process. The experimental nature of the studies presented here meant that we could not properly integrate such considerations into our learning interventions. Below I provide a description of each the variables outlined above and how they might be integrated in a more ecologically valid context.

3.1. Explanation of what the TE aims to provide

For certain structures, it is possible that if learners are given limited support in what to pay attention to in the input, their attentional processes may nevertheless still not be directed to the target feature, even if there are abundant examples of it in the text or if they are visually enhanced. In the context of our study, given the experimental constraints of the study, we did not provide learners with any type of explanation regarding the importance of attending to the textually-enhanced captions. It might, therefore, be beneficial for the learner, if prior to the presentation of TE-captioned media, the instructors would provide an explanation regarding the aims of the TE for the particular structure in question. For example, instructors might highlight

.

the typical problems encountered by L2 learners in their processing of the structure, and how focusing on particular features in the input might facilitate their understanding.

3.2. The provision of more guided instruction

Learners' attention to input, in some cases, may need to be guided and explicit explanation might need to be provided to boost learning gains. Particularly in classroom contexts where the time that can be spent on teaching a particular linguistic construction is limited, explicit explanation of the form and meaning of syntactic structures may speed up the process of language development. Thus, to increase the chances of learning through TE-captioned media, it might additionally be beneficial for instructors not only to include explicit instruction, such as that presented in our short grammar lesson videos, but also, to integrate a more interactive approach. For instance, by providing instructional scaffolding, i.e., the support given to a learner that is tailored to the needs of the student with the intention of helping the student achieve their learning goals (Sawyer, 2006). Such scaffolding can be provided to the learner by highlighting important rules and elaborating on specific examples; by providing clarification or corrective feedback; and by establishing links between what is being learned (Plass & Jones, 2005). In the context of our study, this type instructional scaffolding could have occurred during the presentation of the TE-captioned videos. At certain points, the instructor could have stopped the videos to highlight the usage of a certain form, particularly those that might already be known to be problematic to L2 learners. During these points, instructors could also actively engage with students, clarifying concepts that might not be clear to the learners, providing additional examples, or even establishing connections between current and previously discussed topics. For instance, in Study 3 of this dissertation, the correct usage of the conditional required learners to

.

understand the distinction between the conditional, preterite, imperfect and present tenses. Thus, it might have been advantageous to the learners if the instructors would had been allowed to pause the videos at certain points in order to highlight the similarities and differences between the usage of the conditional tense and the preterite, imperfect and present tenses.

3.3. A more prominent role for the learner

A successful integration of TE-captioned media in the L2 classroom should also include a more prominent role for the learner in their language development. In the context of our studies, several learners reported that the pace at which the videos were presented was too fast for them to process; some mentioned that they would have liked to clarify the meaning of certain words presented in the videos; others mentioned that they would have benefited from taking notes on what they were learning and by being allowed to ask questions about the material presented in the videos. Such interactions are typical in a regular L2 classroom and their integration with techniques such as TE-captioned media would allow the learners to have a more active role in their learning process. In the present studies, for, instance, learners could have been allowed to pause their individual videos to ask questions of clarification; to further attend or process the examples presented in the video at their discretion; and/or to take notes for further study of these examples and thus manage the path and pace of their learning.

In classrooms that support a more digital multimedia environment, the inclusion of additional 'help options' or built-in applications designed to provide assistance, such as "vocabulary annotations, transcripts, translations, control functions (e.g. pause and rewind), listening advice and feedback, dictionaries, or cultural notes" (Cross, 2017), might additionally support a more active and individualized learning environment for the learner. Such techniques are commonly used in computer-assisted language learning (CALL) environments "to draw

.

attention to specific linguistic features and ease the demands of second language processing" (Cárdenas-Claros & Gruba, 2009, p. 70). The integration of such techniques with the TE-captioned videos presented in this dissertation could serve to facilitate learner comprehension of the more complex semantic and syntactic aspects of linguistic input by allowing them to interact with the material itself, for instance, by clicking on problematic words and receiving additional information from multimedia glossaries; by reviewing specific syntactic structures as they view their corresponding examples in the input; and by having access to L1 translations of the input that would allow them to understand how the form is used in their native language.

Both the use of these options and the degree to which they can aid the learner are eminently researchable, particularly at a time when multiple L2 learning platforms are available to millions of users around the world (e.g., Babbel, Duolingo, Yabla). We are beginning to see large-scale interdisciplinary investigations of relevant factors in extensive and varied populations of learners using on-line instruction (Alexopoulou, Michel, Murakami & Meurers, 2017; MacWhinney, 2017; Ziegler et al., 2017). It is only through such large-scale studies that we can have the power to investigate learner by structure by CALL support. Clearly there is promise for supportive multimedia grammar instruction, but we have a long and exciting road ahead in optimizing these advantages.

.

3. References

Alexopoulou, T., Michel M., Murakami A., & Meurers, D. (2017). Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing natural language processing techniques. *Language Learning, 67*(S1), 180-208.

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*(2), 197-235.

Cárdenas-Claros, M. S. & Gruba, P. A. (2009). Help options in CALL: A systematic review. *CALICO Journal* 27(1), 69-90.

Cross, J. (2018). Help options for L2 listening in CALL: A research agenda. *Language Teaching, 50*(4), 544-560.

Han, Z., Park, E. S., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics, 29*(4), 597-618.

Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions. *Studies in Second Language Acquisition, 38*(3), 401-430.

Lee, M., & Révész, A. (2018). Promoting grammatical development through textually enhanced captions: An eye-tracking study. *The Modern Language Journal, 102*(3), 557-577.

MacWhinney, B. (2017). A shared platform for studying second language acquisition. *Language Learning, 67*(S1), 254-275.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41*(3), 720-739.

Muñoz, C. (2017). The role of age and proficiency in subtitle reading. An eye-tracking study. *System, 67*, 77-86.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417.

Winke, P. M. (2013). The effects of input enhancement on grammar learning and comprehension. *Studies in Second Language Acquisition, 35*(2), 323-352.

.

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.

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.

Sawyer, R. K. (2006). Introduction: The new learning science. In R.K. Sawyer. (Ed.). *The Cambridge Handbook of The Learning Science*. Cambridge University Press.

Simard, D., & Foucambert, D. (2013). Observing noticing while reading in L2. In J. M. Bergsleithner, S.N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 207–226). Honolulu, HI: National Foreign Language Resource Center, University of Hawai'i at Mānoa.

Ziegler, N., Meurers, D., Rebuschat, P., Ruiz, S., Moreno-Vega, J. L., Chinkina, M., Li, W., & Grey, S. (2017). Interdisciplinary research at the intersection of CALL, NLP, and SLA: Methodological implications from an input enhancement project. *Language Learning, 67*(S1), 210-232.