This is the Pre-Published Version.
Accepted version [19-Jul-2022]

Testing the Competing Systems Hypothesis: Further evidence from aspect in tutored L2 Spanish

Tiffany Judy*
Wake Forest University

Eloi Puig-Mayenco
King's College, London

Adel Chaouch-Orozco
Hong Kong Polytechnic University

Fernando Martín-Villena
Universidad de Granada

David Miller
University of Illinois, Chicago

*Corresponding author: judyte@wfu.edu

Abstract

This study tests the Competing Systems Hypothesis (CSH; Long & Rothman, 2013; Rothman, 2008) as applied to adult second language acquisition of aspect in Spanish. The CSH purports that differences among tutored and untutored learners (e.g., Goodin-Mayeda & Rothman, 2007; Robison, 1990; Salaberry, 2002) result from competition between one system of underlying grammatical knowledge and another of learned metalinguistic knowledge in TLs. Twenty-nine L1 English-L2 Spanish tutored learners (TLs) and 29 native Spanish speakers completed three tasks examining knowledge of Spanish aspect. A Cloze Task targeted typical use of preterit and imperfect morphology, while a Semantic Entailments Task assessed acquired knowledge of entailments falling out from convergence on the aspectual system. Finally, a Binary-Choice Task targeted three pedagogical simplifications common to L1 English-L2 Spanish textbooks: adverbials that 'go with' the preterit or imperfect, 'meaning-changing' verbs in preterit, and the lower frequency of stative verbs in preterit. Results align with the CSH in that TLs performed above 80% accuracy on preterit and imperfect items in the Cloze Task, indicating acquisition of Spanish aspectual morphology. The Semantic Entailments Task further shows TLs converge on underlying knowledge of grammatical aspect by making a significant distinction between logical and illogical entailments. Finally, competition between the underlying grammatical system and learned metalinguistic knowledge is evidenced in the Binary-Choice Task where TLs are less accurate with preterit; specifically, TLs demonstrate effects from pedagogical simplifications regarding adverbials and 'meaning-changing' verbs.

Keywords: aspect, semantic entailments, Spanish, tutored learners, Competing Systems Hypothesis

This is the accepted version of the publication Judy, T., Puig-Mayenco, E., Chaouch-Orozco, A., Martín-Villena, F., & Miller, D. (2023). Testing the Competing Systems Hypothesis: Further evidence from aspect in tutored L1-English–L2-Spanish. Second Language Research, 39(4), 1165-1189. DOI: 10.1177/02676583221123954

1. Introduction

This study tests the Competing Systems Hypothesis (CSH; Long & Rothman, 2013; Rothman, 2008) by investigating second language (L2) acquisition of aspect and the effects of explicit instruction in L1 English-L2 Spanish tutored learners (TL) in the USA. Aspect gives temporal information about the internal structure of verbs, such as the boundedness of an event (Arche, 2006; Comrie, 1976; Smith, 1991) and it can be encoded in the lexical class of verbs (i.e., lexical aspect) or grammaticalized (i.e., grammatical aspect) with dedicated morphology on the verb itself:

(1a) Leo preparó el café. Atelic activity verb, Perfective aspect
Leo prepare.PRET the coffee
'Leo prepared the coffee.'

(1b) Leo preparaba el café. Atelic activity verb, Imperfective aspect
Leo prepare.IMP the coffee
'Leo was preparing[1] the coffee.'

The acquisition of aspectual features in non-native grammars has been extensively investigated (for overviews see Bardovi-Harlig, 2000; Kempchinsky & Slabakova, 2005; Li & Shirai, 2000; Salaberry, 2008; Salaberry & Shirai, 2002; Slabakova, 2001). Early studies testing the Lexical Aspect Hypothesis (LAH; Andersen, 1986, 1991; Andersen & Shirai, 1994) revealed differing levels of support for the LAH depending on the type of learner (tutored or untutored see Bardovi-Harlig & Bergström, 1996; Bergström, 1995; Hasbún, 1995; Ramsey, 1990; Robison, 1990; Salaberry, 1998, 1999, 2002). Pedagogical intervention affects the input, both type and quantity, to which each type of learner is exposed, a topic we return to below.

A series of studies by Montrul and Slabakova examined semantic knowledge as a proxy for acquisition of the underlying grammatical representation of tense and aspect in Spanish. The intermediate and advanced L1 English-L2 Spanish TLs of Slabakova and Montrul (2002) showed morphological knowledge of preterit/imperfect, as well as relevant semantic entailments (boundedness vs. unboundedness) with accomplishment (e.g., dismantle), achievement (e.g., arrive), and stative (e.g., know) verbs. Under a similar methodological design, Montrul and Slabakova further demonstrated that TLs have knowledge of preterit and imperfect morphology before demonstrating knowledge of their semantic entailments. Montrul and Slabakova (2003)

[1] The gerund form is one gloss of the Spanish imperfect, which can have a continuous, habitual, progressive, or intentional meaning (see Domínguez et al. 2011; Salaberry, 2011).

extended the purview of their original study by including advanced, superior, and near-native L1 English-L2 Spanish speakers, who demonstrated that acquisition of the Spanish aspectual system is possible. These studies were initially couched within generative questions concerning the critical period (Lenneberg, 1967; Penfield & Roberts, 1959), transfer and access to UG (Schwartz & Sprouse, 1994, 1996), and the possibility of acquiring new functional features (Hawkins & Chan, 1997). Together, they provide robust evidence that adult second language learners can acquire the underlying representation of tense and aspect in Spanish despite representational differences between Spanish and English. Thus, Montrul and Slabakova's work highlights the benefit of examining comprehension, particularly, semantic entailments, to reveal the underlying system.

An unresolved question at the forefront of second language acquisition (SLA) research, however, is how to account for the observable variability in the performance of L2 learners. Taking the stance that production may underestimate or overestimate learners' knowledge (Goad et al., 2003; Lardiere, 2000, 2009; Prévost & White, 2000), subsequent formal approaches assume that while learners may show variable knowledge of morphology, interpretation may still be intact as a product of convergence on the underlying representation. Counted among proposals that claim that surface-level production may obscure otherwise target-like L2 acquisition (e.g., Haznedar & Schwartz, 1997; Lardiere, 2000, 2009; Prévost & White, 1999, 2000; Slabakova, 2008, 2013, 2016), is the CSH (Long & Rothman, 2013; Rothman, 2008). Similar to observations raised by Salaberry (2002), the CSH maintains that classroom instruction impacts L2 development. Specifically, the CSH claims that TLs develop two L2 systems that compete at the level of performance. The first system is comprised of underlying grammatical knowledge, making it akin to the grammatical system untutored learners have. The second system is comprised of learned metalinguistic knowledge obtained via classroom instruction. Importantly, TLs are, presumably, exposed to both explicit pedagogical instruction and naturalistic input by way of interactions with native speakers, study abroad, travels, etc. The CSH hypothesizes that, throughout language development, the learner draws on both systems, specifically during performance. Divergence between the two systems is claimed to influence the actual competence of the tutored L2 learner, particularly when competence is measured by performance alone. Even for near-native proficiency speakers, residual effects from explicit instruction may surface during performance tasks.

Goodin-Mayeda and Rothman (2007) later examined advanced L1 English-L2 Spanish and L1 English-L2 Portuguese learners' knowledge of accidental/non-accidental entailments

associated with the preterit/imperfect in Spanish and Portuguese. Since all the L2 participants were university instructors of Spanish or Portuguese, the experiment examined whether instruction on preterit/imperfect influences even highly proficient participants. A semantic task examining accidental/non-accidental entailments revealed that each learner group performed in line with the native groups. The production task targeted use of the preterit with stative verbs, so-called 'meaning-changing' verbs (*no quiso* 'refused' vs. *no quería* 'didn't want to'), and adverbial phrases taught as generally associated with either preterit (*ayer* 'yesterday) or imperfect (*siempre* 'always'). This task revealed that both Portuguese and Spanish participants favored imperfect with statives more than native speakers. 'Meaning-changing' verbs were particularly problematic for the L2 participants. The L2 Portuguese participants did not differ from native speakers for the other verb types (accomplishment, achievement, or activity); however, L2 Spanish speakers differed significantly from native speakers on activity verbs, again preferring imperfect. Rothman and Iverson (2008) investigated accidental/non-accidental semantic entailments as well as group- and kind-denoting readings of the preterit and imperfect in tutored Portuguese L2ers. In line with previous studies examining semantic entailments, the results revealed a developmental pattern with advanced L2ers performing more native-like than intermediate L2ers. In fact, advanced L2ers only differed from the control group in the preterit condition containing adverbial quantifiers (e.g. *sempre* 'always'). While intermediate L2ers were less polarized than the control and advanced participants in their assessments of preterit and imperfect, intra-group comparisons demonstrated that the intermediate group differentiated between the preterit and imperfect. Finally, Rothman (2008) examined both tutored and untutored L2ers to isolate the effects of pedagogy. The 20 near-native tutored L2ers were very familiar with the instruction of preterit and imperfect, each having taken a minimum of five years of Spanish language courses and each teaching Spanish language courses at the time of testing. The 11 near-native untutored L2ers had never received formal instruction in Spanish, but rather had lived in a Spanish-speaking country for a minimum of seven years. Participants completed a contextualized cloze task and a written production task (n=12) targeting the pedagogical simplifications of preterit with statives and adverbials that are taught as being associated with either the preterit (e.g. *ayer* 'yesterday') or imperfect (e.g. *siempre* 'always'). Rothman reported that, whereas untutored L2ers had target-like performance in both tasks, the highly advanced tutored L2ers showed some residual effects of explicit instruction where it differs from actual distribution of preterit and imperfect, specifically with statives in the preterit and with

adverbs like *siempre* 'always'. The results of Rothman (2008) were interpreted as supporting the CSH claim that simplified pedagogical rules concerning the preterit and imperfect can result in target-deviant performance even in highly advanced L2ers (Long & Rothman, 2013).

The current study contributes to the rich body of research examining acquisition of aspect in L1 English-L2 Spanish learners (in addition to aforementioned studies, see, for example Domínguez et al., 2011; Domínguez et al., 2013; Domínguez et al., 2017) by examining data that tap semantic knowledge of aspect. To test the predictions of the CSH, the current study enhances the methodology of Montrul and Slabakova (2003) and Rothman (2008) in three important ways. First, to ensure that each TL had an underlying grammatical system for aspect in Spanish, we included a modified and improved version[2] of the sentence conjunction judgment task in Montrul and Slabakova (2003). Second, we conducted a review of Spanish language textbooks aimed at native speakers of English to substantiate the claims made by early CSH studies that said textbooks employ the pedagogical simplifications tested therein. This review, outlined in section 2.2, validates the ubiquity of these simplifications, lending support to the CSH's claim that Spanish students in the USA are exposed to these simplifications (and naturalistic input). Finally, in addition to including the cloze task employed in Rothman (2008), we expanded the main experimental task that targeted simplified pedagogical rules to include three experimental conditions. While the 12 items in Rothman (2008) included stative verbs and 'meaning-changing' verbs lumped together with some items also displaying adverbials, each condition herein was crafted to isolate each of the three pedagogical simplifications associated with L1 English-L2 Spanish instruction described in section 2.2. The benefit of this design is twofold in that it allows us to determine if each of the pedagogical simplifications are equally problematic for tutored learners and if and when it is possible for them to overcome each simplification. Thus, the data allow us to examine the underlying grammatical knowledge of Spanish preterit/imperfect, as well as performance, across TLs of various proficiencies to determine if the CSH can account for the variability found not only in highly advanced learners as demonstrated by Rothman (2008), but also in earlier stages of L2 development. Finally, this approach bridges SLA and applied linguistics (Marsden, 2018; Slabakova & Marsden, 2019), revealing the potential effects of pedagogical simplifications and areas that pedagogical practices may improve.

[2] The modified version eliminates inconsistencies related to token length, idiomatic phrases, structure, and matches the illogical and logical pairs much more closely such that their appreciable differences rest on aspect.

2. Background

2.1 Aspect in Spanish

Lexical aspect refers to the fundamental makeup of verbal predicates (and their constituents) and is contingent upon the concept of telicity. Telic events are those with an intrinsic endpoint, such as *Owen cooked a nice dinner*, and atelic events are those with no intrinsic endpoint, such as *Owen cooks nice dinners*. At the lexical level, telicity describes the aspectual properties of the verb, referring always to endpoints. Grammatical aspect, by contrast, is expressed morpho-syntactically on the verb itself—in perfective and imperfective forms—where the boundedness of an event refers to real boundaries in time and space. Grammatical aspect is also often referred to as *viewpoint aspect* because it is concerned with situation-internal time (Comrie, 1976; Smith, 1986), where use of one aspect or the other depends on the speaker's view of the event. The perfect form (i.e., the preterit) is bounded in time, having a beginning or end, and is viewed from the outside with no concern for internal structure. The imperfect form is unbounded, having no beginning or end, and is viewed from the inside with specific regard to its internal structure (Depraetere, 1995).

Importantly for this study, Spanish has two morphological forms that denote perfectivity and imperfectivity, and whose distribution is governed by discourse-related factors. These aspectual contrasts are obligatorily marked in past tense, such that past tense morphology indicates both tense and aspect. The preterit encodes past tense and perfective aspect (i.e., bounded event), while the imperfect encodes past tense and imperfective aspect (i.e., unbounded event):

(2a) Juan escribió un libro.
Juan write.PRET a book
'Juan wrote a book.'

(2b) Juan escribía un libro.
Juan write.IMP a book
'Juan was writing a book.'

Example (2b) shows that both the Spanish morphology *–ía* and the English equivalent in the progressive indicate an event underway. While there is a potential endpoint to the event of writing a book, thus being telic, imperfective morphology does not indicate a specific endpoint or beginning and is, therefore, unbounded.

Notably, Spanish has a reliable morphological contrast between the completed or uncompleted status of events, whereas English uses the same form for habitual events in the preterit and imperfect. Generally, although the difference in tense and aspect morphology can interact with the aspectual value of predicates, atelic events tend to occur with the imperfect tense, whereas telic events tend to occur with the preterit. Notwithstanding, given that the use of preterit or imperfect is not arbitrary, but rather governed by aspect, the preterit and imperfect can occur with any predicate in Spanish, depending on discourse-related factors such as the speaker's intention. The examples below illustrate this point:

(3) El partido fue/era divertido. Atelic
The game be.PRET /be.IMP fun
'The game was fun.'

(4) Carlos escuchó/escuchaba música. Atelic
Carlos listen.PRET/listen.IMP music
'Carlos listened/was listening to music.'

(5) Ari compró/compraba un libro. Telic
Ari buy.PRET/buy.IMP a book
'Ari bought/was buying a book.'

(6) Carolina descubrió/descubría la verdad. Telic
Carolina discover.PRET/discover.IMP the truth
'Carolina discovered/was discovering the truth.'

In sum, and importantly for acquisition of Spanish aspect, we assume that adult learners bring a fully-developed L1 grammar that may differ from the L2 grammar to the acquisition task (Schwartz & Sprouse, 1994, 1996). As shown above, English lacks the morphological paradigm that distinguishes between perfective and imperfective aspect that Spanish employs. Such grammatical differences between languages are known to affect the acquisition process, where learners must learn new rules and unlearn old ones (see Bardovi-Harlig, 2000; Comajoan, 2014; Dietrich et al., 1995; Domínguez et al., 2013; Montrul, 2004; Salaberry, 2008).

With this in mind, an adult L1 English-L2 learner of Spanish must learn that Spanish has two morphological paradigms to encode aspect while English has only one. Subsequently, learners must map perfective and imperfective aspect onto unique targets in Spanish. In addition to the linguistic factors that affect this process, below we acknowledge and explore the effects that instruction may have.

2.2 Aspect in tutored L2 Spanish

Early aspect studies pointed to learner type (tutored vs. untutored) effects. Salaberry and Shirai (2002) report that textbooks and classroom activities at the time focused on preterit forms. For example, Hasbún (1995) for Spanish and Bergström (1995) for French revealed that the preterit is frequently, if not exclusively, taught before the imperfect, making it a likely candidate for a default 'past tense'. Furthermore, Kaplan (1987) and Swain (1992) showed that French classroom activities favor the preterit, making this form more frequent and salient in learners' primary linguistic data.

To validate the claims made in the CSH regarding pedagogical simplifications and where these may diverge from actual language use in Spanish, the chapters/sections on preterit and imperfect from 15 Spanish language textbooks designed for native English-speaking students were analyzed to determine if previous findings hold today. The review included five beginner, seven intermediate, and three advanced textbooks published by well-known publishing houses between 2003 and 2020. Four relevant components were examined:

(a) presentation of preterit before imperfect,

(b) explicit association of certain adverbials with either preterit or imperfect,

(c) explicit mention of 'meaning-changing' verbs, and

(d) occurrence of stative verbs in preterit and imperfect

The presence and, where appropriate, frequency of each component was carefully tallied. Before addressing the findings, we note that our participants have varying experience with the specific textbooks reviewed below. However, covering textbooks from different publishers and proficiency levels allows us to demonstrate that the same patterns are present across many textbooks. Thus, it is not the case that only beginners are exposed to these simplifications if they use one of the five beginner textbooks examined herein, but rather that at some point in their studies, they will be taught these simplifications, if not repeatedly. Importantly for the methodology of the study, the outcomes of this review directly inform the design of Task 3 (see section 3.2.3). It also bears stating explicitly that an additional aim of this article is to connect SLA and applied linguistics such that they may be mutually informative (Marsden, 2018; Slabakova & Marsden, 2019). Therefore, the purpose of this section is not to criticize these 15 textbooks for

presenting pedagogical simplifications, particularly because their approach to the presentation of the preterit and imperfect is not unique. Rather, they are presented herein as the standard and analyzed to determine the extent to which these pedagogical simplifications are present.

Similar to previous reviews, a commonality across all five beginner textbooks is that the preterit is presented first in explanatory passages, with both a description of form and function. Likewise, six of the seven intermediate textbooks present the preterit before introducing the imperfect; here, the seventh textbook, like two of the advanced textbooks, does not include a description or rules of formation or of use for either the preterit or imperfect. Finally, the third advanced textbook presents a review of both aspects concurrently, starting with the imperfect.

Table 1. Presentation of preterit before imperfect

	Presentation of preterit before imperfect
Beginner	5/5
Intermediate	6/7
Advanced	0/3

Since the preterit is generally described as representing a completed action, time-bounding adverbials such as *ayer* 'yesterday', and *una vez* 'one time' frequently appear in textbook practice exercises/activities. Importantly, these adverbials are employed in textbooks that have not yet presented the imperfect, thus increasing the likelihood that students make early associations between certain adverbials and the preterit. Four of the five beginner textbooks examined include adverbials in preterit practice exercises. The fifth textbook provides a list of adverbials typically associated with the preterit. With varying frequency, the intermediate textbooks also include time-bounding adverbials in preterit practice exercises. These are largely absent, though not entirely so, in the advanced textbooks.

Similarly, a component of instruction common to beginner textbooks is the explicit mention of, by way of a demarcated box or specific section, adverbials that typically appear with the imperfect, such as *siempre* 'always' and *mientras* 'while'. Four of the five beginner textbooks provide such a list, and the same number of textbooks employ these adverbials in practice exercises explicitly requiring use of imperfect. The use of adverbials in conjunction with the imperfect is less pronounced in the intermediate textbooks (four of seven), and only one intermediate textbook includes an explicit list of adverbials typically used with the imperfect. Finally, only one of three

advanced textbooks provides consistent examples of adverbials said to be typically associated with the imperfect in exercises calling for the imperfect, and none include an explicit list.

Table 2. Explicit association of certain adverbials with either preterit or imperfect

	Preterit with time-bounding adverbs	Imperfect with non-time-bounding adverbs
Beginner	4/5	4/5
Intermediate	6/7	4/7
Advanced	1/3	1/3

In Spanish language textbooks geared towards English speakers, a common, and often prominent, component of instruction is the so-called 'meaning-changing' verbs. This group of verbs consists of verbs that are presented as having different meanings in English. For example, *conocías* 'you knew' is contrasted with *conociste* 'you met'. All five beginner textbooks examined include a section on the 'meaning-changing' verbs[3], and this instruction is repeated in five of the seven intermediate and one of the three advanced textbooks where it appears in the instructor version only as a note to remind students. Importantly, one advanced textbook specifically states that the common instruction that the verbs 'change meaning' is inaccurate.

Table 3. Explicit mention of 'meaning-changing' verbs

	'Meaning-changing' verbs
Beginner	5/5
Intermediate	5/7
Advanced	1/3

Lastly, the distribution of statives in preterit and imperfect across all 15 textbooks was tallied. In four of five beginner textbooks, more statives appeared in the imperfect practice exercises than in the preterit ones. This difference reached up to 42.8% more, while there were only 3.5% more statives in the preterit than in the imperfect in the one textbook falling outside the norm. In the simultaneous presentation of the two forms, all five introductory textbooks contained more statives in the imperfect exercises than in the preterit ones (range 6.2-53% more). Two of the seven intermediate textbooks do not present practice exercises for the imperfect; of the five that present simultaneous preterit/imperfect practice exercises, more statives appear in the imperfect practice exercises than in the preterit ones (range 8.2-54.4%). Similarly, two of the textbooks do not present simultaneous preterit/imperfect practice exercises. All five that do contain more

[3] One does not refer to them specifically as 'meaning-changing' verbs.

statives in the imperfect exercises than in the preterit ones (range 14.3-67.3% more). Of the three advanced textbooks examined, one does not have practice exercises related to the preterit and imperfect[4]; the remaining two include more statives in the imperfect practice exercises than in the preterit ones (range 19.5-71.4%). Only one of the three advanced textbooks presents activities to practice with both forms simultaneously; this textbook includes more statives in the imperfect than the preterit as well (24.9%).

Table 4. Higher frequency of stative verbs in the imperfect

	More statives in imperfect than preterit
Beginner	4/5
Intermediate	5/7
Advanced	2/3

Thus, the first area where pedagogical instruction simplifies the distribution of preterit and imperfect regards the adverbials students are exposed to with both the preterit and imperfect, especially at the beginner level. Given the explicit mention and frequency of use of adverbials in exercises that already instruct students to use one form or the other, students may rely on the presence of adverbs like *ayer* 'yesterday' or *siempre* 'always' to determine or understand the aspect of a sentence. The picture that emerges from the Spanish language textbook presentations and practice activities examined is that certain adverbs 'go with' certain forms. Contrary to these generalizations, some adverbials that are explicitly taught as being associated with one form can, and do, modify events in the other form. Explicit focus on the adverbials as overt clues that reveal the aspect of the verb may shift attention away from the intended meaning of the sentence/verb, and give students the false impression that the adverbials themselves determine aspect. Secondly, regarding the 'meaning-changing' verbs, insinuating that you 'meet' someone in the preterit but you 'know' someone in the imperfect shifts learners' focus from the (un)boundedness of the verb itself and towards a (sometimes erroneous) English-based translation of Spanish aspect. For example, if a learner wishes to express the bounded statement of 'I knew her my whole childhood, (but we lost touch)', explicit instruction regarding *conocer* 'to know' in preterit and imperfect would push the learner to erroneously produce *la conocía toda mi infancia*, largely because it is odd to be meeting someone for your entire childhood. Thus, while this simplification may initially

[4] Unlike the other two advanced textbooks which contain explanations, practice exercises and writing passages, this textbook focuses exclusively on the mechanics composition writing.

help students to understand the aspectual difference between 'to know' and 'to meet' in English, its continued repetition may have detrimental effects in that it encourages learners to use English translation and it does not make clear that both forms may be used with both English meanings. Thirdly, the distribution of statives occurring in imperfect exposes learners to Spanish input that seems to inadvertently amplify the role of lexical aspect, whereby telic verbs are more frequent in the preterit and atelic verbs are more frequent in the imperfect. While it is unlikely that learners have metalinguistic knowledge of lexical or viewpoint aspect, they can certainly understand that verbs in the preterit are completed while verbs in the imperfect are continuous, habitual, progressive, or intentional, which is the common description in textbooks. Thus, when confronted with an atelic verb like *ser* 'to be', learners may be reticent to use this verb in the preterit due to its stative nature, thus relying on lexical aspect as opposed to the boundedness of the event that they wish to describe, especially given the infrequency of this verb in the preterit in the input.

3. Study and design

Based on the claims of the CSH and textbook review above, our study entertains the following research question:

- Is TL performance regarding Spanish aspect affected by the presence of two competing grammatical systems across development?

The CSH hypothesizes that TLs will show variability in L2 performance across development and claims that this variability affects even highly advanced TLs. Specifically, TLs are hypothesized to have underlying grammatical knowledge of aspectual in Spanish as evidenced by their sensitivity to semantic entailments (Task 2) (Montrul & Slabakova, 2003; Slabakova & Montrul, 2003). Similarly, where instruction aligns with actual distribution of preterit/imperfect, TLs are predicted to demonstrate target-like knowledge of aspect (Task 1 and Task 2). However, TLs of all proficiencies are predicted to show effects of instruction where instruction does not match actual use of preterit/imperfect (Task 3). Thus, we also ask:

- If TLs show effects of instruction at the performance level (Task 3), do these effects manifest equally across the three targeted pedagogical simplifications?

Following the claims of the CSH, there is no *a priori* expectation that TLs' performance will be differentially affected by each pedagogical simplification. Thus, we hypothesize that they will be equally affected.

3.1 Participants

In total, 29 native Spanish speakers (NSS) and 29 tutored L2 learners of Spanish (TLs) participated in this study.[5] The NSS group is comprised of speakers of Peninsular Spanish and the TL group is comprised of TLs of Spanish in U.S. institutions. All participants completed three background tasks, including a Spanish proficiency test, an English proficiency task, and the *Bilingual Language Profile* (BLP; Birdsong et al., 2012). The proficiency task is an adaptation of the *DELE* (2002) containing 50 items across two sections: a multiple-choice vocabulary section (n = 30) and a cloze section (n = 20). Finally, the BLP assessed language dominance among bilinguals via 19 questions concerning language attitudes, history, proficiency, and use. Since immersion has been shown in some studies to be highly variable (DeKeyser, 2007), information regarding self-reported months of immersion in a Spanish-speaking country was also gathered; nonetheless, initial statistical analyses explored the effect of length of immersion , and indicated that immersion was not a significant predictor. As such, this factor was not further explored. Table 5 presents participant demographics.[6]

Table 5. Participant demographics

	NSS (n=29)	TLs (n=29)
Self-reported gender	15 (female)	16 (female)
Average age (SD)	33.59 (14.17)	24.14 (4.12)
Average Spanish proficiency (SD)	47.86 (1.70)	39.14 (1.70)
Average BLP score (SD)	125.19 (47.30)	-111.36 (29.07)
Average Length of Immersion in Months (SD)	n/a	15.72 (10.08)

NSS=Native Spanish Speakers; TL=Tutored L2 Learners; BLP=Bilingual Language Profile

[5] Eighty-four participants (29 NSS and 55 TLs) completed the experiment. However, results from Tasks 1 and 2 served as exclusion criteria to ensure acquisition of typical uses of preterit/imperfect in Spanish and target-like aspectual configuration. In Task 1, participants failing to meet the 75% accuracy cut-off were excluded from the final analysis (n=6). A two-pronged approach was applied to Task 2. First, participants whose mean ratings for logical items that were below 2.5 (the midpoint in the task's 4-point Likert scale) or above 2.5 for illogical items were excluded (n=13). Second, participants whose mean ratings to logical vs. illogical items were not significantly different were excluded (n=7). As a final safeguard for Task 3, the task of interest, any participant whose accuracy to the distractor items was below 75% were excluded (n=0). Thus, of the original 55 TLs, 29 met the inclusion criteria.

[6] This study was conducted in accordance with the recommendations of and approved by the author's research ethics committees. All participants gave informed consent in accordance with the Declaration of Helsinki.

3.2 Experimental Tasks

In addition to the background tasks, participants completed three untimed experimental tasks in Spanish. The first task consisted of a cloze task version of Goldilocks and the Three Bears used in Rothman (2008). The second task was a semantic entailment task adapted from Montrul and Slabakova (2003). The third task was a cloze task targeting mismatches between commonly used textbook explanations and naturalistic use of preterit/imperfect in Spanish.[7]

3.2.1 Task 1: Cloze Task

As in Rothman (2008), participants saw a story containing 55 missing past tense verbs requiring preterit or imperfect and were instructed to select the form that best fit each blank. Importantly, each item contained instances of preterit and imperfect that correspond to canonical uses common to the L2 classroom where the teaching of these contexts matches the usage. Items in this task did not contain 'meaning-changing' verbs or adverbials, two of the contexts tested in Task 3. Responses were coded in a binary fashion capturing target selection of preterit/imperfect. Example (7) below shows an imperfect and preterit example, respectively, with target responses bolded.

(7) Imperfect and preterit token
Como ya (fue/**era**) mediodía, los osos **(se sentaron**/se sentaban) a comer.
'As it was already midday, the bears sat down to eat'

3.2.2 Task 2: Semantic Entailments Task

The second task was a semantic entailments task adapted from Montrul and Slabakova (2003), meant to capture underlying knowledge of grammatical aspect. Participants rated the perceived logicalness of individual sentences on a scale from 1 (*completamente ilógica* 'completely illogical') to 4 (*completamente lógica* 'completely logical'). The task contained 42 experimental and 14 distractor items. The experimental items were equally divided across three conditions: achievement verbs (n = 14), statives (n = 14), and accomplishments (n = 14).[8] Half the items for each condition, including distractors, were logical (n = 7) and half were illogical (n = 7) . Responses were coded on a 4-point scale.

[7] The materials for the three tasks will be archived on the IRIS database (www.irisdatabase.org) upon acceptance of the manuscript.
[8] Following Montrul and Slabakova (2003), activity verbs were not included herein because they produce contradictory statements predicated on logic as opposed to aspect, the focus of the experiment.

(8) Achievements
Los González vendían/*vendieron la casa pero nadie la compró.
'The González family was selling/*sold their house but no one bought it.'

(9) Statives
La película era/*fue a las 7 pero empezó a las 7:30.
'The movie was (supposed to be) at 7 but started at 7:30.'

(10) Accomplishments
Amanda llevaba/*llevó el paquete hasta la oficina pero lo perdió.
'Amanda was taking the package to the office but lost it.'

3.2.3 Task 3: Binary-Choice Task

Finally, the third task was a binary-choice task targeting uses of preterit and imperfect that contradict common L2 Spanish classroom practices. The review in section 2.2 directly informs the design of this task in that it validates that claim made by the CSH that the three simplifications regarding preterit and imperfect use are common to Spanish textbooks for native English-speakers in the USA. First, recall that textbooks instruct that certain adverbials 'go with' the preterit (*ayer* 'yesterday') or the imperfect (*siempre* 'always'). Second, the review demonstrated that certain verbs (*saber* 'to know') are presented as having different meanings in preterit ('to find out') as compared to imperfect ('to know'). Third, statives appear more frequently in imperfect. Although these pedagogical strategies are meant to aid tutored learners, they do not reflect the full reality of usage preterit and imperfect use in Spanish. Thus, in this task, participants saw 36 experimental items that directly contradicted these generalizations and 18 distractor items. The experimental items were equally divided across three conditions: adverbials (n = 12), the so-called 'meaning-changing' verbs (n = 12), and statives (n = 12).[9] Half the items for each condition targeted the preterit (n = 6) by bounding the timeframe of the event and half targeted the imperfect (n = 6). Half of the distractors required preterit (n = 9) and half required imperfect (n = 9). Responses were coded in a binary fashion capturing target selection of preterit and target selection of imperfect.

Imperfect
(11) Adverbials

[9] Three items were removed due to less than 50% accuracy in NNS responses (two items from the Adverbials condition, with 6.90% and 37.93% accuracy, and one item from the Meaning-changing condition, with 41.38% accuracy). A fourth item was removed from the Meaning-changing Preterit condition because it confounded two experimental variables.

Marcos siempre (hizo/**hacía**) los deberes que le mandaba la maestra.
'Marcos always did the tasks that his teacher gave him.'

(12) 'Meaning-changing' verb
Adriana (quiso/**quería**) ir al cine los sábados.
'Adriana wanted to go to the cinema on Saturdays.'

(13) Statives
Mi vecino me dijo que (hubo/**había**) muchos perros en el parque.
'My neighbor told me that there were many dogs in the park.'

Preterit

(14) Adverbials
El testigo siempre (**dijo**/decía) la verdad mientras duró el juicio.
'The witness always said the truth while the trial lasted.'

(15) 'Meaning-changing' verbs
Valentina (**pudo**/podía) ser la campeona en su mejor época.
'Valentina managed to be the champion in her best days.'

(16) Statives
(**Hubo**/Había) varias manifestaciones durante los años 60.
'There were many protests during the 1960s.'

4. Results

The statistical analyses were performed using R ((version 3.3.1) (R Core Team, 2016)), with the *lme4* package (Bates et al., 2015). A generalized linear mixed-effects model with a binomial family fitted to the accuracy data (Baayen, 2008; Baayen et al., 2008), and linear mixed-effects on the continuous data from the ratings in Task 2 were completed. For model selection, we followed the procedure recommended by Gries (2018). The error data in Tasks 1 and 2 were dummy-coded as 1 (correct) and 0 (incorrect). First, we started with a maximal structure for fixed-effects in a model that also included a basic random structure (i.e., intercepts for items and participants). Then, backward selection was employed, extracting the factor or interaction that explained the least amount of variance from the maximal model. Next, we performed a maximum likelihood ratio comparison between both models. The model offering a better fit to the data was selected to continue the process until having a model of best fit. Additionally, a maximal random structure was included by adding random slopes for items and participants by Tense. Following Matuschek et al. (2017), alternative models with simplified random structures were used. These were then compared to the maximal model—again, we used maximum likelihood ratios where 0.2 was the

significance level—to see if they improved the fit. All the simplified models offered a better fit than the maximal one.

4.1 Task 1: Cloze Task

Table 6 shows Task 1 accuracy scores for both groups in the Imperfect and Preterit conditions.

Table 6. Task 1 Accuracy percentages (standard deviations)

	Imperfect	Preterit
NSS	95.93% (19.79)	91.01% (28.61)
TLs	82.76% (37.80)	88.61% (31.79)

NSS=Native Spanish Speakers; TLs=Tutored L2 Learners

The following factors were included in the data analysis: Tense (preterit or imperfect), Group (NSS and TLs), Proficiency, and Dominance; the latter two factors were modelled as continuous variables quantified by the scores from the DELE and the BLP. The final model of best fit had the following fixed effects and interactions: Tense, Group, Proficiency, Dominance, and the interaction of Tense*Dominance (see Appendix A for final model and its outcome). As random effects, the model had intercepts for items and participants. The results showed a significant main effect of Group ($\beta = 2.37$, $p < .001$), Tense ($\beta = 1.27$, $p < .05$), Proficiency ($\beta = 0.32$, $p < .001$), and a significant interaction between Tense*Dominance ($\beta = 0.66$, $p < .001$). The main effect of Group indicates that, overall, NSS are more accurate than TLs in their responses. The main effect of Tense shows that participants are more accurate with items in the Preterit condition. The effect of Proficiency ($\beta = 0.32$, $p < .001$) reflects that the more proficient the participants are in Spanish, the more accurate their responses are. Lastly, the significant interaction between Tense and Dominance indicates that, in the Imperfect condition, participants who are less dominant in Spanish were less accurate than more dominant ones. In the Preterit condition, the differences in Spanish dominance do not modulate accuracy.

4.2 Task 2: Semantic Entailments Task

Table 7 shows Task 2 results for both groups, Logical and Illogical conditions divided by VerbClass (accomplishment, achievement, stative). Recall that the lowest possible score is 1 (completely illogical) and the highest is 4 (completely logical).

Table 7. Task 2 Ratings (standard deviations)

	Logical Condition			Illogical Condition		
	ACC	**ACH**	**STA**	**ACC**	**ACH**	**STA**
NSS	3.16 (1.14)	3.60 (0.82)	3.82 (0.49)	1.75 (1.05)	1.33 (0.75)	1.72 (1.09)
TLs	3.06 (1.20)	3.06 (1.20)	3.56 (0.84)	1.87 (1.11)	1.49 (0.87)	1.92 (1.11)

NSS=Native Spanish Speakers; TLs=Tutored L2 Learners; ACC=Accomplishment; ACH=Achievement; STA=Stative

The factors considered in the model were Group, Logicalness, VerbClass, Proficiency and Dominance. The model of best fit included fixed effects of Group, Logicalness, VerbClass, Proficiency and Dominance, and the interactions of Logicalness*VerbClass and Logicalness*Dominance. As random effects, the model contained random intercepts for items and participants, and a random slope for participants by Logicalness (see Appendix B for final model and its outcome). The results showed a main significant effect of Logicalness ($\beta = 1.24$, $p < .001$), and two two-way interactions: Logicalness*VerbClass ($\beta = 0.77$, $p < .05$) and Logicalness*Dominance ($\beta = 0.17$, $p < .05$). The main effect of Logicalness reflects that participants in both groups gave higher ratings to Logical than Illogical items. Regarding the interaction of Logicalness*VerbClass, post-hoc pairwise comparisons were performed employing the *lsmeans* package (version 2.30-0; Lenth, 2016) with familywise error rates and the Tukey method. The comparisons showed that ratings to different verb classes in the Logical condition are significantly different than in the Illogical one (for accomplishment verbs, $\beta = -1.29$, $p < .001$; for achievement verbs, $\beta = -2.07$, $p < .001$; and for stative verbs, $\beta = -1.87$, $p < .001$). The interaction between Logicalness*Dominance indicates that in the Logical condition, the more dominant in Spanish a participant is, the more logical they find the sentences in that condition), whereas in the Illogical condition, the more dominant in Spanish a participant is, the more illogical they find the sentences in that condition.

4.3 Task 3: Binary-Choice Task

Table 8 shows Task 3 results for both groups for Adverbial, Meaning-Changing, and Stative conditions in Imperfect and Preterit.

Table 8. Task 3 Accuracy percentages (standard deviations)

	Imperfect	Preterit

	Adverbial	Meaning-changing	Stative	Adverbial	Meaning-changing	Stative
NSS	96.55% (18.30)	97.13% (16.75)	97.70% (15.03)	69.83% (46.10)	83.62% (37.17)	97.70% (15.03)
TLs	91.34% (28.15)	90.81% (28.98)	85.06% (35.75)	28.45% (45.31)	25.86% (43.98)	77.59% (41.82)

NSS=Native Spanish Speakers; TLs=Tutored L2 Learners

The model of best fit included the following fixed effects: Group, Tense, Condition, Proficiency, Dominance, as well as several interactions (see Appendix C for final model and its outcome). The results showed significant interactions of Group*Tense (β = 1.19, p < .01), Condition*Dominance (β = -0.56, p < .05), and Condition*Tense (β = -3.02, p < .001). Post-hoc analyses revealed that accuracy in the Preterit is significantly higher in NSS than in TLs (β = -2.46, p < .001). Additionally, Imperfect responses are not significantly different across the three conditions. Importantly, accuracy differs significantly between Tenses in the Adverbial (β = 3.07, p < .001) and Meaning-changing conditions (β = 3.02, p < .001), but not in the Stative condition.

To explore the hypothesis that TL s have greater difficulty with the preterit as compared to imperfect due to classroom instruction, we conducted post-hoc pairwise comparisons. NSS results showed that accuracy in the three conditions is not significantly different across the two forms. However, TL responses to Adverbial (β = 3.66, p < .001) and Meaning-changing conditions (β = 3.62, p < .001) are significantly less accurate in Preterit than in Imperfect, while the difference across the Stative condition is not significant. Lastly, TL s are significantly less accurate than NSS with Adverbial (β = -2.46, p < .01) and Meaning-changing verbs (β = -2.46, p < .01) in the Preterit.

5. Discussion

To test the Competing Systems Hypothesis (Rothman, 2008; Long & Rothman, 2013) , we entertained the following two research question:

- Is tutored L2 performance regarding Spanish aspect affected by the presence of two competing grammatical systems across L2 development?
- If TLs show effects of instruction at the performance level (T ask 3), do these effects manifest equally across the three targeted pedagogical simplifications?

For Task 1 (Cloze Task), the prediction that TLs would be highly accurate in the typical contexts of preterit/imperfect was borne out. The statistical analyses showed significant effects of Tense and Proficiency, as well as a significant interaction of Tense*Dominance. Proficiency modulated accuracy, whereby the more advanced participants were, the better they performed in both preterit and imperfect. The significant interaction showed that the more dominant participants were in Spanish, the better they performed on imperfect, whereas irrespective of dominance, TLs did well on preterit. Recall that a commonality across all five beginner textbooks reviewed is that preterit form and function are presented first. Furthermore, studies have shown that classroom activities favor the preterit (Kaplan, 1987; Swain, 1992), making this form more frequent and salient in the learners' primary linguistic data.[10] Interestingly, the results show that typical uses of imperfect are slightly more problematic than those in the preterit condition. This may be explained by the order of presentation and amount of exposure to both preterit and imperfect forms in the L2 Spanish classroom.[11] Furthermore, empirical work on the acquisition of L2 Spanish aspect has shown that specific uses of imperfect are more problematic than those of the preterit (e.g., Domínguez et al., 2011). Still, while the statistical analysis confirms that the imperfect is more problematic, both participant groups had mean accuracy scores above 80%, indicating that the bound/unbound nature of the preterit and imperfect has been acquired. Thus, it seems likely that these TLs have already overcome any initial 'default past tense' for preterit that they may have had at lower proficiency levels.

For Task 2 (Semantic Entailments Task), TLs that have acquired the bound/unbound nature of preterit/imperfect were hypothesized to show sensitivity to semantic entailment violations arising from aspect. Results revealed that both NSS and TLs made a significant distinction between logical and illogical semantic entailments as indicated by the effect of Logicalness. Importantly, post-hoc comparisons on the interaction of Logicalness*VerbClass revealed that both NSS and TLs treated the logical and illogical items in each verb class significantly differently, indicating that TLs' sensitivity is not isolated to one verb class, but rather generalized. The interaction of

[10] As pointed out by a reviewer, we acknowledge that the input learners are exposed to might be dependent on how the instructors decide to present the two forms. However, from personal experience, we maintain that instructors normally follow the order of presentation found in textbooks, and thus, it is probable that preterit is taught before imperfect.

[11] A reviewer also mentioned that it may be that the imperfect is simply more difficult to process given that it is the source of the aspectual contrast most notable in Spanish for English speakers (Doiz, 2002).

Logicalness*Dominance indicated that the more dominant participants were in Spanish, the more likely they were to rate logical statements higher. This does not mean that participants less dominant in Spanish were not sensitive to the distinctions, though, since they all showed sensitivity to the violations. These results are in line with previous work on L2 Romance languages (Goodin-Mayeda & Rothman, 2007; Montrul & Slabakova, 2003; Rothman & Iverson, 2008; Slabakova & Montrul, 2003). Crucially, combined with evidence from Task 1, we conclude that the TLs have underlying grammatical knowledge of aspect in Spanish.

The question that remains is whether pedagogical simplifications may lead to residual effects at the level of performance. Recall that Task 3 (Binary-Choice Task) tested three pedagogical simplifications common to Spanish language textbooks geared towards English speakers: adverbials that 'go with' the preterit or imperfect, 'meaning-changing' verbs, and the disproportionate frequency of statives in imperfect (see Section 2.2). Because these simplifications do not fully encompass the reality of preterit/imperfect distribution in Spanish, the CSH purports that learners will have two systems of knowledge (one that comprises the underlying grammatical knowledge and another comprising learned metalinguistic knowledge) that will compete at the level of performance. From this claim, TLs of all proficiencies are predicted to show effects of L2 instruction at the level of performance in carefully designed conditions where pedagogical simplifications do not match actual uses of preterit/imperfect. Results from Task 3 showed a significant interaction for Group*Tense, revealing differences in the treatment of preterit items across groups. Specifically, the TLs are significantly less accurate in preterit for all three simplifications. Additionally, the significant interaction for Condition*Tense showed that TLs are not equally accurate across the three conditions (adverbials, 'meaning-changing', statives): contrary to Goodin-Mayeda and Rothman (2007) and Rothman (2008), the TLs more readily overcome the simplifications surrounding statives than those related to 'meaning-changing' verbs or adverbials.

The question, then, is why the TLs are more target-like with statives compared to 'meaning-changing' verbs and adverbials. Examination of the distribution of statives in Section 2.2 revealed that they are more frequently presented in imperfect than preterit regardless of textbook level, which provides little evidence that advanced TLs's accuracy with them is a result of instruction. A post-hoc explanation of our results would be that unlike textbooks, naturalistic data reveals more equal distributions across aspect. For example, the frequency of *haber* 'there

is/are' and *estar*[12] 'to be' in Peninsular Spanish of the *Corpus de Referencia del Español Actual* (CREA) is more balanced: *había* (36.49%) versus *hubo* (29.73%), *estaba* (35.74%) versus *estuvo* (41.47%). This shows that while textbooks lean towards imperfect statives, a more balanced distribution is present in naturalistic input learners are exposed to.[13] Recall that TLs tested herein reported an average of 15 months of immersion in Spanish, thus making it likely that they have had access to naturalistic input where statives would frequently appear in preterit. Still, we must consider why naturalistic input has not mitigated (to the same degree) effects of classroom instruction on adverbials and 'meaning-changing' verbs. A clear difference between the three simplifications examined in Task 3 is that explicit instruction is often provided for adverbials and 'meaning-changing' verbs, but not to the same degree for statives where simple frequency was measured. Recall from Section 2.2 that Spanish textbooks provide lists of adverbials commonly used with each form and reinforce the connection by employing adverbials from the list in exercises that focus exclusively on one form or the other, especially in beginner textbooks. Similarly, explicit instruction regarding 'meaning-changing' verbs prompts learners to use English translation to determine which form to use. Both pedagogical strategies shift the focus away from what truly determines which form is used (i.e., the (un)boundedness of the event) and towards learned cues (adverbials, English translations) that may or may not reflect target-like use of Spanish preterit/imperfect. Thus, in line with the CSH, competition between the systems of underlying grammatical knowledge and learned metalinguistic knowledge, especially with a more developed pedagogical awareness concerning adverbials and 'meaning-changing' verbs in the preterit/imperfect, may offer a tenable explanation for the improved performance of TLs with statives as compared to adverbials and 'meaning-changing' verbs.

6. Conclusion

To conclude, this study enriches our understanding of L2 acquisition of preterit/imperfect and the effects of pedagogical instruction. Experienced instructors and researchers examining convergence in this grammatical domain alike are aware of the difficulties L1 English-L2 Spanish leaners face with preterit and imperfect. While many studies report differences between native Spanish

[12] These verbs were chosen as examples because they were used in the experimental tasks.

[13] We acknowledge that this post-hoc explanation of our data needs to be further investigated.

speakers and learners (for overviews see Bardovi-Harlig, 2000; Comajoan, 2014; Dietrich et al., 1995; Salaberry, 2008), studies focusing on learners' semantic knowledge, presumed to result from underlying competence, report target-like sensitivity especially at higher proficiency levels (Goodin-Mayeda & Rothman, 2007; Montrul & Slabakova, 2002, 2003; Rothman & Iverson, 2008; Slabakova & Montrul, 2003). Early studies by Montrul and Slabakova reveal the importance of examining comprehension; likewise, the careful design of the tasks used herein improves upon similar previous studies. Specifically, the inclusion of two measures of acquisition allowed us to establish that the TLs converged on the bounded/unbounded distinction in Spanish preterit/imperfect, respectively (Task 1 and 2 results). Task 1 also serves as a reference point when comparing outcomes where task design matches (Task 1) or mismatches (Task 3) instruction. Furthermore, Task 3 improved on its predecessor (Rothman, 2008) by separating statives and 'meaning-changing' verbs (which are stative) into distinct categories to individually assess the effect of their associated pedagogical simplifications, validated by the textbook review. While Rothman (2008) reports residual effects of instruction with adverbials and statives in the preterit, the TLs examined herein only showed effects for adverbials and 'meaning-changing' verbs but are more accurate on statives. Thus, teasing these variables apart has provided greater granularity regarding TLs' performance and potential effects of instruction.

Still, in line with previous work on the CSH, the results ultimately validate the claim that TLs develop two L2 systems that compete at the level of performance only. Results from Task 1 and 2 establish that TLs distinguish between boundedness in Spanish while results from Task 3 point to the influence of instruction. While semantic entailments such as those presented in Task 2 are not explicitly taught, the typical uses of preterit and imperfect in Task 1 and their more nuanced uses in Task 3 are. Comparing across Task 1 and 3, results revealed, that where instruction matches distribution, both systems converge and TLs perform target-like; however, where simplifications do not encompass available uses, TLs' acquired and metalinguistic systems compete, resulting in non-target-like performance on adverbials and 'meaning-changing' verbs. Future replications should ideally include an untutored L2 group.

A final point to consider is what these results and the claims of the CSH mean for foreign language teaching. As demonstrated, pedagogical simplifications can result in unintended effects in performance even in highly advanced learners. We do not intend to suggest that instruction should eliminate these simplifications altogether, especially since they capture a good portion of

the distribution of preterit/imperfect and provide a good starting point for learners. Nonetheless, what is meant to initially aid students in grasping a notoriously difficult grammatical concept might have undesirable effects precisely where distribution of preterit/imperfect falls outside these general rules. As the textbook review revealed, these simplifications are present, to varying degrees, across textbooks produced by different publishers and aimed at different proficiencies, meaning that the simplifications are likely reinforced as students' proficiency increases. If the input learners receive is not supplemented with naturalistic preterit and imperfect examples like those presented in Task 3, the system of learned metalinguistic knowledge may overshadow the one with the underlying grammatical knowledge. Thus, learners may be best served by a gradual phase-out of these simplifications, with more emphasis on nuanced uses of preterit/imperfect, like those seen in Task 3, as proficiency increases. Finally, as pointed out by Marsden (2018) and Marsden and Slabakova (2019), researchers in both acquisition and applied fields should strive to reduce the gap between research and pedagogy, as a bi-directional bridge between these two communities can ultimately benefit the L2 learning experience and outcomes.

References

Andersen, R. (1986). El desarrollo de la morfología verbal en el español como segundo idioma. In J. M. Meisel (Ed.), *Adquisición de lenguaje/Aquisição da linguagem* (pp. 115–138). Vervuert.

Andersen, R. (1991). Developmental sequences: The emergence of aspect marking in second language acquisition. In T. Huebner & C. Ferguson (Eds.), *Crosscurrents in SLA and linguistic theories* (pp. 305–324.). John Benjamins.

Andersen, R., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition, 16*(2)*,* 133–156. https://doi.org/10.1017/S0272263100012845

Arche, M. J. (2006). *Individuals in time: Tense, aspect and the individual/stage distinction*. John Benjamins.

Ayoun, D., & Salaberry, R. (Eds.) (2005). *Tense and aspect in Romance languages: Theoretical and applied perspectives*. John Benjamins.

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

Bardovi-Harlig, K., & Bergström, A. (1996). The acquisition of tense and aspect in SLA and FLL: A study of learner narratives in English (SL) and French (FL). *Canadian Modern Language Review*, *52*, 308–330.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bergström, A. (1995). *The expression of past temporal reference by English-speaking learners of French* [Unpublished doctoral dissertation]. Pennsylvania State University, State College PA.

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). Bilingual Language Profile: An easy-to-use instrument to assess bilingualism. *COERLL, University of Texas at Austin*. <https://sites.la.utexas.edu/bilingual/>

Comajoan, Ll. (2014). Tense and aspect in second language Spanish. In K. L. Geeslin (Ed.), *The handbook of Spanish second language acquisition* (pp. 235–252). Blackwell.

Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems* (Vol. 2). Cambridge University Press.

DeKeyser, R. (2007). Study abroad as foreign language practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from Applied Linguistics and Cognitive Psychology* (pp. 208–226). Cambridge.

Depraetere, I. (1995). On the necessity of distinguishing between (un)boundedness and (a)telicity. *Linguistics and Philosophy*, *18*(1), 1–19. https://doi.org/10.1007/BF00984959

Dietrich, R., Klein, W., & Noyau, C. (1995). *The Acquisition of temporality in a second language*. John Benjamins.

Doiz, A. (2002). The preterit and the imperfect as grounding predications. In F. Brisard (Ed.), *Grounding: The Epistemic Footing of Deixis and Reference* (pp. 299–347). Mouton de Gruyter.

Domínguez, L., Arche, M. J., & Myles, F. (2011). Testing the predictions of the Feature-Assembly Hypothesis: Evidence from the L2 acquisition of Spanish aspect morphology. In N. Danis, K. Mesh, & H. Sung (Eds.), *BUCLD 35: Proceedings of the 38th Annual Boston University Conference on Language Development* (pp. 185-196). Cascadilla Press.

Domínguez, L., Arche, M. J., & Myles, F. (2017). Spanish imperfect revisited: exploring L1 influence in the reassembly of imperfective features onto new L2 forms. *Second Language Research*, *33*(4), 431–457. https://doi.org/10.1177/0267658317701991

Domínguez, L., Tracey-Ventura, N., Arche, M. J., & Myles, F. (2013). The role of dynamic contrasts in the L2 acquisition of Spanish past tense morphology. *Bilingualism: Language and Cognition*, *16*(3), 558–577. https://doi.org/10.1017/S1366728912000363

Goad, H., White, L., & Steele, J. (2003). Missing inflection in L2 acquisition: Defective syntax or L1-constrained prosodic representations? *Canadian Journal of Linguistics*, 48(3-4), 243–263. https://doi.org/10.1017/S0008413100000669

Gries, S. (2018). Course: Statistics for Linguistics with R [course materials]. University of Reading: https://research.reading.ac.uk/celm/statistics-for-linguists-with-r-5-days-duration/

Hasbún, L. (1995). *The role of lexical aspect in the acquisition of tense and grammatical aspect in Spanish as a foreign language* [Unpublished doctoral dissertation]. Indiana University.

Hawkins, R., & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition: The 'Failed Functional Features Hypothesis'. *Second Language Research*, *13*(3), 187-226. https://doi.org/10.1191/026765897671476153

Haznedar, B. & Schwartz, B. D. (1997). Are there optional infinitives in child L2 acquisition? In E. Hughes, M. Hughes & A. Greenhill (Eds.), *Proceedings of the 21st Boston University Conference on Language Development* (pp. 257–268). Cascadilla.

Kaplan, M. (1987). Developmental patterns of past tense acquisition among foreign language learners of French. In B. VanPatten, T. Dvorak, & J. F. Lee (Eds.), *Foreign language learning: A research perspective* (pp. 52–60). Newbury House.

Kempchinsky, P., & Slabakova, R. (Eds.) (2005). *Aspectual inquiries*. *Studies in Natural Language and Linguistic Theory*; No. 62. Springer.

Lafford, B. (1996, October 24–27). *The development of tense-aspect relations in L2 Spanish narratives: Evidence to test competing theories* [Paper presentation]. Second Language Research Forum 96, Tucson, Arizona. http://web.archive.org/web/19990128183542/http:/www.coh.arizona.edu:80/slrf96/

Lardiere, D. (2000). Mapping features to forms in second language acquisition. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 102–129). Blackwell.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research*, *25*(2), 173–227. https://doi.org/10.1177/0267658308100283

Lenneberg, E. H. (1967). *Biological foundations of language.* Wiley.

Lenth, R. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. https://doi:http://dx.doi.org/10.18637/jss.v069.i01

Li, P., & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect*. Mouton de Gruyter.

Long, D., & Rothman, J. (2013). Generative approaches and the Competing Systems Hypothesis: Formal acquisition to practical application. In J. Schweiter (Ed.), *Theoretical and pedagogical innovations in SLA and bilingualism* (pp. 63-83). John Benjamins.

Marsden, H. (2018). Generative second language acquisition and language teaching: Advancing the dialogue. In J. Cho, M. Iverson, T. Judy, T. Leal, & E. Shimanskaya (Eds.), *Meaning and structure in second language acquisition: In honor of Roumyana Slabakova* (pp. 283–308). John Benjamins.

Marsden, H., & Slabakova, R. (2019). Grammatical meaning and the second language classroom: Introduction. *Language Teaching Research*, *23*(2), 147–157. https://doi.org/10.1177/1362168817752718

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, R. H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Montrul, S. (2004). Psycholinguistic evidence for split intransitivity in Spanish L2. *Applied Psycholinguistics*, *25*(2), 239–267. https://doi.org/10.1017/S0142716404001122

Montrul, S., & Slabakova, R. (2002). The L2 acquisition of morphosyntactic and semantic properties of the aspectual tenses preterite and imperfect. In A. T. Pérez-Leroux & J. Muñoz-Liceras (Eds.), *The acquisition of Spanish morphosyntax (*pp. 115–151). Kluwer.

Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers. *Studies in Second Language Acquisition, 25*(3)*,* 351–398. https://doi.org/10.1017/S0272263103000159

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms.* Princeton University Press.

Prévost, P., & White, L. (2000). Missing Surface Inflection or impairment in second language acquisition? *Second Language Research, 16*(2), 103–133. https://doi.org/10.1191/026765800677556046

Ramsey, V. (1990). *Developmental stages in the acquisition of the perfective and the imperfective aspects by classroom L2 learners of Spanish* [Unpublished doctoral dissertation]. University of Oregon.

REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [online]. *Corpus de referencia del español actual.* <http://www.rae.es> [12/06/2020]

Robison, R. (1990). The primacy of aspect: Aspectual marking in English interlanguage. *Studies in Second Language Acquisition*, *12*(3), 315–330. https://doi.org/10.1017/S0272263100009190

Rothman, J. (2008). Aspect selection in adult L2 Spanish and the Competing Systems Hypothesis: When pedagogical and linguistic rules conflict. *Languages in Contrast, 8*(1), 74–106. https://doi.org/10.1075/lic.8.1.05rot

Rothman, J., & Iverson, M. (2008). Poverty-of-the-stimulus and L2 epistemology: Considering L2 knowledge of aspectual phrasal semantics. *Language Acquisition: A Journal of Developmental Linguistics*, *15*(4), 270–314. https://doi.org/10.1080/10489220802352206

Salaberry, R. (1998). The development of aspectual distinctions in classroom L2 French. *Canadian Modern Language Review*, *54*(4), 508–542.

Salaberry, R. (1999). The development of past tense verbal morphology in classroom L2 Spanish. *Applied Linguistics*, *20*(2), 151–178.

Salaberry, R. (2000). *The development of past tense morphology in L2 Spanish*. John Benjamins.

Salaberry, R. (2002). Tense and aspect in the selection of past tense verbal morphology. In R. Salaberry & Y. Shirai (Eds.) *Tense-Aspect morphology in L2 acquisition* (pp. 397–415). John Benjamins.

Salaberry, R. (2008). *Marking past tense in second language acquisition: A theoretical model*. Continuum.

Salaberry, R. (2011). Assessing the effect of lexical aspect and grounding on the acquisition of L2 Spanish past tense morphology among L1 English speakers. *Bilingualism: Language and Cognition*, *14*(2), 184–202. https://doi.org/10.1017/S1366728910000052

Salaberry, R., Comajoan, L., & Gonzalez Gonzalez, P. (2013). Research findings and pedagogical practices on the teaching of L2 tense-aspect. In R. Salaberry & Ll. Comajoan (Eds.), *Research design and methodology in studies on second language tense and aspect* (pp. 423–450). Mouton de Gruyter.

Salaberry, R., & Shirai, Y. (2002). *Tense-Aspect morphology in L2 acquisition*. John Benjamins.

Schwartz, B., & Sprouse, R. (1994). Word order and nominative case in nonnative language acquisition: a longitudinal study of (L1 Turkish) German interlanguage. In T. Hoekstra & B. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar* (pp. 317–368). John Benjamins.

Schwartz, B., & Sprouse, R. (1996). L2 cognitive states and the Full Transfer/Full Access Model. *Second Language Research*, *12*(1), 40–72. https://www.jstor.org/stable/43104504

Slabakova, R. (2001). *Telicity in the second language.* John Benjamins.

Slabakova, R. (2008). *Meaning in the second language*. Mouton de Gruyter.

Slabakova, R. (2013). What is easy and what is hard in second language acquisition: A generative perspective. In M. D. P. García Mayo, M. J. Gutierrez-Mangado, & M. Martínez Adrián (Eds.), *Contemporary Approaches to Second Language Acquisition (AILA Applied Linguistics Series)* (pp. 5–28). John Benjamins.

Slabakova, R. (2016). *Second language acquisition*. Oxford University Press.

Slabakova, R., & Montrul, S. (2003). Genericity and aspect in L2 acquisition. *Language Acquisition*, *11*(3), 165–196. https://doi.org/10.1207/s15327817la1103_2

Smith, C. (1986). A speaker-based approach to aspect. *Linguistics and Philosophy*, *9*(1), 97–115. https://www.jstor.org/stable/25001234

Smith, C. (1991). *The Parameter of aspect*. Kluwer.

Swain, M. (1992). Manipulating and complementing content teaching to maximize learning. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood-Smith, & M. Swain (Eds.), *Foreign/Second Language Pedagogy Research* (pp. 234–250). Multilingual Matters.

Appendix A: Final Statistical model outcome of Task 1

	Coefficient	SE	z-value	p-value
Intercept	1.65	.33	5.04	< .001
Group (NSs)	2.37	.66	3.58	< .001
Tense	1.27	.55	2.3	< .05
Proficiency	.32	.16	1.99	< .05
Dominance	-0.44	.34	-1.31	.19
Tense (Preterit) by Dominance	-0.66	.21	-3.17	< 0.01

Code:

Accuracy ~
Tense + Condition + Spanish Proficiency + Dominance +
Tense: Dominance +
(1 + Tense | Subject) + (1 + Tense | Item)

Appendix B: Final Statistical model outcome of Task 2

	Coefficient	SE	t-value	p-value
Intercept	1.98	.25	8.1	< .001
Group (TLs)	-0.19	.21	-0.9	.37
Logicalness	1.24	.26	4.82	< .001
VerbClass (Achievement)	-0.4	.25	-1.63	.11
VerbClass (Stative)	.01	.26	.03	.98
Proficiency	-0.05	.06	-0.88	.38
Dominance	-0.15	.11	-1.34	.18
Logicalness (Logical) by VerbClass (Achievement)	.77	.34	2.31	< 0.05
Logicalness (Logical) by VerbClass (Stative)	.58	.34	1.71	.1
Logicalness (Logical) by Dominance	.17	.07	2.65	< 0.05

Code:

```
Response ~
        Group + Logicalness + VerbClass + Proficiency + Dominance +
        Logicalness : VerbClass
        + Logicalness : Dominance +
        (1 + Logicalness | Subject) + (1 | Item)
```

Appendix C: Final Statistical model outcome of Task 3

	Coefficient	SE	z-value	p-value
Intercept	2.2	.45	4.88	< .001
Group (NSs)	1.28	.66	1.94	.05
Dominance	.15	.33	.44	.66
Proficiency	.28	.15	1.9	.06
Tense (Preterit)	-0.64	.54	-1.19	.23
Condition (Adverbial)	.2	.56	.35	.72
Condition (Change)	.33	.57	.59	.55
Group (NSs) by Tense (Preterit)	1.19	.39	3.1	< .01
Dominance by Condition (Adverbial)	-0.56	.22	-2.54	< .05
Dominance by Condition (Meaning-change)	.02	0.24	0.1	.92
Tense (Preterit) by Condition (Adverbial)	-3.02	0.8	-3.78	< .001
Tense (Preterit) by Condition (Meaning-change)	-2.97	0.81	-3.69	< .001

Code:

```
Accuracy ~
        Group + Dominance + Proficiency + Tense + Condition  +
        Group : Tense +  Dominance : Condition + Tense : Condition +
        (1 | Subject) + (1 | Item)
```