CrossMark

Applied Psycholinguistics **38** (2017), 477–508
doi:10.1017/S0142716416000308

Processing the Spanish imperfect subjunctive: Depth of processing under different instructional conditions

SERGIO ADRADA-RAFAEL
Fairfield University

Received: July 7, 2015 Accepted for publication: May 29, 2016

ADDRESS FOR CORRESPONDENCE
Sergio Adrada-Rafael, Department of Modern Languages and Literatures, Canisius Hall, Room 215, Fairfield University, Fairfield, CT 06824. E-mail: sadradarafael@fairfield.edu

ABSTRACT
This study aimed to examine 88 intermediate learners of second language Spanish' depth of processing under different types of instruction when being exposed to a complex grammar form, the imperfect subjunctive, on a reading task. Depth of processing was measured by employing think-aloud protocols. Due to the absence of reactivity, six silent and think-aloud conditions were merged into three experimental ones differing in explicitness. The study followed a pre-, post-, and delayed test design with 2 weeks between tests, which consisted of interpretation, controlled-production, and comprehension tests. Overall, results showed that participants in the more explicit condition produced more instances of processing, and that deeper processing correlated with a more accurate production of the target form and with a higher comprehension of the reading passage.

In the field of second language acquisition (SLA), numerous studies have examined the effects that different types of instruction varying in explicitness have on L2 learning (e.g., Benati, 2004; Erlam, 2003, 2005; Farley, 2004; Herron & Tomasello, 1992; Robinson, 1996; Rosa & Leow, 2004a; Rosa & O'Neill, 1999; Sanz, Lin, Lado, Stafford, & Bowden, 2014; Shaffer, 1989; Stafford, Bowden, & Sanz, 2012; Tagarelli, Borges Mota, & Rebuschat, 2014), yielding inconclusive findings. While most of these studies have looked at the "product" of the treatment, the final result; the "process," that is, what is going on in learners' minds while carrying out the tasks, has been widely overlooked. The internal processes taking place in the learner's mind while concurrently completing a specific task can shed some light on how deeply the target form(s) have been processed, and how that correlates with the actual results found in the study. Understanding depth of processing (DOP) was addressed in the field of cognitive psychology in the 1970s, referred to as levels of processing (Craik, 2002; Craik & Lockhart, 1972; Craik & Tulving, 1975), but

© Cambridge University Press 2016 0142-7164/16

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

has been more recently examined in the field of SLA by employing online and/or offline techniques (Hsieh, Moreno, & Leow, 2015; Hulstijn, 1997; Laufer & Hulstijn, 2001; Leow, Hsieh, & Moreno, 2008; Morgan-Short, Heil, Botero-Moriarty, & Ebert, 2012). However, while this "depth" has been categorized by the amount of cognitive effort put into the task and by the noticing of a target item or items (for a review of studies addressing DOP, see Leow & Mercer, 2015), it has not been examined under more or less explicit types of instruction. Because the role of formal language instruction is seen as key to promote language acquisition (e.g., Doughty, 1991), and instruction might differ in terms of how it is delivered to learners, addressing DOP in relation to different instructional styles is crucial to understand how different instructional types might induce a deeper or a shallower processing. The goal of the present study is to address and fill this gap in the literature, examining participants' DOP, via think-aloud protocols, while learning a complex grammar form, the Spanish imperfect subjunctive, under three different instructional conditions differing in explicitness (more explicit, less explicit, and baseline).

REVIEW OF THE LITERATURE

DOP in cognitive psychology

Craik and Lockhart (1972) coined the term DOP in the field of cognitive psychology; it refers to the idea that rather than memory stores, we attend to and process information in a continuum. This continuum ranges from shallow sensory analyses to deeper analyses concerned with meaning and implication. Craik and Lockhart also argued that deeper levels of processing were associated with a longer lasting effect of the acquired item in memory. This scheme could apply to verbal materials, such as phonological analyses, word identification, or sentence comprehension. Processing the meaning of a new lexical item would take place at a deeper level than processing its phonological form, which would occur at a more shallow level. Craik and Tulving (1975) further explored these notions in a series of experiments. Participants were presented words and were asked questions varying in their degree of complexity. Positive responses indicated the DOP of a participant with regard to a specific word. They found empirical evidence for the effects of levels of processing, with deeper processing leading to greater retention in terms of recognition and recall ability. While facing some criticism in the following years (Baddeley, 1978; Nelson, 1977) concerning the lack of an independent index of depth, it has remained an influential model of processing, and further empirical studies investigated the effects of these levels of processing with different types of items and modes, such as L1 words in the auditory mode (Karayianni & Gardiner, 2003; Rajaram, 1993; Toth, 1996), L1 words in the visual mode (Gardiner, Java, & Richardson-Klavehn, 1996; Java, 1994; Rajaram, 1993; Richardson-Klavehn & Gardiner, 1995), or in pictures (Gardiner, Brandt, Vargha-Khadem, Baddeley, & Mishkin, 2006; Gardiner, Gregg, Mashru, & Thaman, 2001; Konstantinou & Gardiner, 2005). The main conclusion that can be extracted from the previous studies is that processing at a shallow level is sufficient to encode the new items in the semantic memory, and that processing of items at a deeper level registers them in the episodic memory. Craik (2002), in a revision of his

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

theory of levels of processing and empirical studies related to it, acknowledged that the operationalizations of DOP had not been very satisfactory and that they had lacked an objective index of DOP, addressing Nelson (1977) and Baddeley's (1978) main criticism to his theory. As explained in the next section, DOP has been reinterpreted in the SLA field as degree of elaboration, more in line with Craik and Tulving's study than with the original semantic versus perceptual distinction proposed by Craik and Lockhart.

Empirical studies in the field of SLA

In the field of SLA, adult second language (L2) learners are known to have a limited processing capability (McLaughlin, 1987). Learners are exposed to input in written or aural form, and part of this input will actually be processed by the learner and become intake. How much of the input becomes intake will depend on a variety of factors (e.g., content complexity and length, grammar complexity, and pragmatic content). VanPatten (1994) reviews the three different processes involved in successful L2 development. The first one is *input-to-intake*, where some structures of the input are noticed by the learner; the second one is *intake-to-acquisition*,[1] where some of those noticed structures move one step further and become integrated as part of the learner's repertoire; and a third and final process is *acquisition-to-use*, where the L2 learner makes use of that repertoire to communicate (cf. VanPatten & Cadierno, 1993). VanPatten and Jegerski (2011) pointed out more recently that research in processing has been using online and offline measures, with online measures being increasingly employed in the field. Learners may look identical when looking at the product, "but may have arrived at the interpretation in different ways. Only online methods can allow us to see these subtle differences in processing" (p. 5). They also distinguished between two different types of processing: processing as part of the acquisition of formal features (i.e., how learners create a linguistic system) and processing that assumes or presupposes underlying grammatical knowledge during comprehension (i.e., how learners put it into practice).

In the attentional strand of research, Shook (1994), in an attempt to probe deeper into the input-to-intake stage, examined the processing of two target forms in Spanish: the present perfect tense and the relative pronouns *que*, *quien(es)*. The author provided 125 participants with reading passages with the target forms embedded. Participants were first- or second-year students of Spanish. The input was modified and presented under the following conditions: no alteration (control), target form enhancement, and target form enhancement with focus on grammatical rules. Assessment tasks included postexposure recognition and production tests for both present perfect and relative pronouns. Overall, results yielded a significant improvement when attention had been drawn to the target forms (enhancement conditions). In general, second-year students outperformed first-year students in present perfect production, with an opposite trend observed for relative pronouns production. On the recognition assessment tests, both years performed better on the present perfect form than on relative pronouns. Shook argued that the recognition results could be expected because the present perfect form is more meaningful than the relative pronouns. However, this argumentation would fail to interpret

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

production results, where first-year students outperformed second-year ones on relative pronouns. In addition, Shook argued that further language experience would be expected to translate into a better performance on either target form, which is not the case for relative pronoun production. The author offered no plausible explanation for this finding. One key limitation in the interpretation of these findings is that we do not know if learners paid the appropriate amount of attention to the targeted items in the text, as Leow (2015) pointed out. Similarly, Gass, Svetics, and Lemelin (2003) observed that the way Shook manipulated the input did not ensure that learners would focus their attention on something. Therefore, it is important to keep in mind that Shook's attempt to address DOP as amount of attention needs to be interpreted with caution.

Qi and Lapkin (2001) investigated the role of processing in an L2 writing task with its subsequent feedback. They conducted a case study with two Mandarin speakers who were English as a second language students with different levels of proficiency. Participants completed a written task in three different stages. First, they wrote a story based on a writing prompt provided by the researchers who then corrected the composition, reformulating it and making it sound more nativelike. Second, 4 days later, both participants received the reformulated version and the original one and had to think-aloud while comparing both versions. Third, 3 days after Stage 2, both learners had to write a new version, incorporating the modifications they had received at Stage 2. Analyses of the think-alouds to account for instances of processing revealed that deeper ones resulted in a higher number of corrections, which led to a better rewrite. With these findings in hand, Qi and Lapkin suggested that learners who showed a deeper processing of the corrections provided by the researchers seemed to yield an overall better performance at Stage 3. The authors concluded that reformulation proved to be a valid pedagogical tool, more than the traditional error correction technique. The implications derived from what think-aloud protocols reveal are also discussed, pointing out the prevalence of quality over quantity of noticing and its relation with the level of proficiency in the L2. The researchers, however, asked for caution when interpreting the results, as no strong claims can be made with a pool of two participants. As they put it, their results constitute "tentative rather than definite answers" (p. 294).

Building on Craik and Lockhart's (1972) *DOP* and on Craik and Tulving's (1975) notion of *elaboration*, Laufer and Hulstijn (2001) developed the *involvement load hypothesis* for L2 vocabulary learning. It consisted of three components: need, search, and evaluation. Need, a more motivational component, was classified as moderate (when imposed by an external agent) or strong (when it was intrinsically motivated). Search and evaluation could also be more or less prominent. Search refers to the learner's attempt to find the meaning of an unknown L2 word by different means (using a dictionary, etc.). However, if the word was provided by the instructor or glossed in the text, there would be no search involved. Evaluation, by contrast, entails the comparison of a given word with other words. The evaluation would be moderate if this comparison took place in a given sentence, and strong if it happened in an original context, where it has to be created. The combination of the three and their degrees of prominence constitute the involvement load engaged in processing the new words. The greater the involvement load, the better the retention. According to Laufer and Hulstijn, a greater involvement load would be equivalent to a deeper level of processing. Their hypothesis was confirmed in

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

subsequent studies (e.g., Hulstijn & Laufer, 2001; Keating, 2008; Laufer, 2003). However, Keating (2008) found that when equating time on task across tasks, the difference on learners' cognitive load was significantly reduced. While Keating treated time as a separate variable, Hulstijn and Laufer (2001) regarded time as an inherent property of the task, allowing more time for the task that involved a higher cognitive load. Therefore, not isolating time as a variable could bring up a potential confound when interpreting the results.

Recently, Bird (2012) conducted a replication of one of Craik and Tulving's (1975) experiments. Participants in Craik and Tulving's study (1975) were native speakers of English exposed to words in English. Bird (2012), however, compared effects of shallow and deep encoding tasks on recognition of target lexical items with nonnative ($n = 24$) and native speakers of English ($n = 24$). Because the nonnative English-speaking participants had Arabic as their first language (L1), Bird posited that their L1 would be of little help when performing the tasks and subsequent tests in English compared to the native speakers. One hundred and twenty target words were split into three blocks of 40 words each, of which 20 were higher frequency and 20 were lower frequency words. Participants were assigned to each of these lists. All words rotated so that they appeared in all four of the following conditions: nonsemantic *yes* and *no* and semantic *yes* and *no*. Results replicated those found by Craik and Tulving (1975) in their original study. Deeper processing led to better recognition than shallow processing for both low and high frequency words, and this applied to both native and nonnative speakers. As in Gass et al. (2003), DOP was predetermined by the inherent characteristics of each condition, and no online measures were employed to see if participants were processing in the expected way. In addition, as with Hulstijn and Laufer (2001), these findings are limited to L2 vocabulary, not including other grammatical forms.

Building on a strand of research that examines the allocation of attention to both form and meaning in the L2 (Greenslade, Bouden, & Sanz, 1999; VanPatten, 1990; Wong, 2001), Leow et al. (2008) employed think-aloud protocols to determine whether participants were processing for both meaning and form. Their study was an extension of VanPatten (1990), who hypothesized that paying attention to grammatical forms would negatively affect participants' processing for meaning. The participants in VanPatten's study were 202 learners of Spanish from different levels: first-semester, fourth-semester, and third-year university students. They were exposed to a 275-word text in Spanish about inflation in Latin America, and the modality of presentation was aural. Learners were randomly assigned to three experimental conditions or a control condition. In the three experimental conditions, subjects had to pay attention to the lexical item *inflación*, the article *la*, or the third-person plural verbal morpheme *–n.* In the control condition, learners only listened for content. Testing for meaning was conducted through a free recall task. Overall, results indicated that the condition attending to the lexical item performed similarly to the control condition, while conditions attending to the article *la* or the verb ending *–n* performed more poorly than the control and lexical conditions. Some differences were found when comparing proficiency levels, as with the ending *–n*. In summary, these findings confirmed his hypothesis that grammatical forms would block processing for meaning.

Leow et al. (2008) fine-tuned VanPatten's methodology in several ways. First, they presented the passage in the written mode, as in Greenslade et al. (1999).

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Second, they replaced the comprehension free recall task by a 10-item multiple-choice test, explaining that potential memory decay could have influenced the low number of recalls found in VanPatten (1990). Third, they also replaced the lexical item *inflación* by the word *sol*, arguing that the former is a cognate of the English term *inflation* and, because it has three syllables, it was more salient than the monosyllabic grammatical items. Fourth, they included a direct object clitic *lo*, suggesting that it is "a form that carries both meaning and grammatical function in the input" (p. 674), and not just a grammatical function. Seventy-two second-semester students of Spanish participants were randomly assigned to five experimental conditions, which included a control group. However, the main methodological addition for the interests of this study was the inclusion of non-metacognitive think-aloud protocols to operationalize the process of reading for meaning, which was their baseline to compare the experimental conditions. The passage was 358 words long, and the participants were 99 second-semester students of Spanish. Findings contradicted those in VanPatten (1990) and Greenslade et al. (1999), and revealed no significant differences in comprehension between conditions, which the authors reasoned could be due to the different modality of the passage: aural versus written. Regarding the think-aloud data, three levels of processing were established, and deeper processing was observed for the lexical item *sol* (73%), than for *la* (45%), *lo*, or *–n* (31%). However, these findings did not report a relationship between processing and comprehension scores given that one limitation of this study, which was noted by the authors, was the small number of participants, which made it difficult to make any strong statistical inference.

In an attempt to further explore Leow et al.'s (2008) findings, Morgan-Short, Heil, et al. (2012) conducted a conceptual replication. In their study, they included the following modifications: (a) a non-think-aloud group was added to address the potential of reactivity while reading for form and meaning, and (b) the sample size was increased to be able to make stronger claims with regard to the relationship between comprehension and DOP. Participants were 361 third-semester students of Spanish. Results revealed a significant effect regarding reactivity but with a minimal effect size, which prevented the authors from making any strong claim on this issue. As in Leow et al. (2008), no significant effect was found between condition and performance in comprehension. Finally, a significant relationship was found between processing levels and performance. The deeper the processing levels, the higher the score on the comprehension test. The authors concluded that a trend where attention to form and meaning does not affect comprehension can be detected in the written modality (Leow et al., 2008). In addition, a correlation between levels of processing and performance emerged with a larger sample size. One of the limitations of this study, also found in Leow et al., was the issue of backtracking during treatment, which was not controlled for. Therefore, some participants could have revisited the reading passage while others did not, creating a potential confound on the interpretation of results. As the authors noted, a computer-administered study could prevent this problem.

Hsieh et al. (2015) recently revisited Hsieh (2008), which compared two types of instructional conditions: a computerized face-to-face instruction that was teacher centered and that provided grammatical explanation with no feedback (i.e., more explicit), and a computer-assisted instructional condition that was learner centered,

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

providing learner-initiated practice with implicit feedback but did not provide any metalinguistic information. Hsieh reported that while similar performances were observed by both groups on the immediate posttest, the less explicit group performed substantially better than the more explicit group on the delayed posttest (2 weeks later). To address this discrepancy in performances, the authors coded the unreported think-aloud protocols in relation to levels of awareness and DOP of the target form: *gustar* (to please) produced by the beginner-level of Spanish participants. Their findings indicated that while the more explicit type of instruction prompted more instances of awareness at the level of understanding, the less explicit condition revealed more instances of deeper processing, showing a higher cognitive effort by participants in this condition during the experimental instructional exposure. The authors argued that, when not provided with metalinguistic information prior to task, learners need to make an additional cognitive effort to try to infer the target form and that this DOP appears to lead to better retention of the target form or structure, as evidenced on the delayed posttests.

MOTIVATION FOR THE PRESENT STUDY

As seen in the above studies, there has not been a consistent pattern of results in the L2 DOP literature. Both Hulstijn and Laufer (2001) and Bird (2012) found that deeper processing was related to better recognition and retention of L2 vocabulary, but both of these studies *assigned* participants to specific processing groups and did not collect an online measure of processing to determine whether participants were processing at the expected levels. In addition, both studies were conducted only at the word level. Shook (1994) found an overall effect when drawing learners' attention to the enhanced forms but also failed to collect online data on learners' processing during treatment. Morgan-Short, Heil, et al. (2012) incorporated think-alouds to account for the relation of DOP and performance more accurately, and found that deeper levels of processing resulted in better performance. Leow et al.'s (2008) findings were in line with those in Morgan-Short, Heil, et al., but did not perform statistical analyses due to the low number of subjects included in the final sample, which limited their generalization.

These results are indicative of the important role that DOP might play in SLA, employing different techniques (online and offline) to measure it. As seen in the latter strand of studies reviewed (Hsieh et al., 2015; Leow et al., 2008; Morgan-Short, Heil, et al., 2012), one of these are verbal protocols. Despite some inherent limitations, such as learners not verbalizing every line of thought, or the potential issue of reactivity, this online measure is useful to categorize how L2 learners are processing the information. With regard to the relationship between types of instruction and DOP**,** this is an area that has yet to be investigated, and would be important in explaining the outcomes found for L2 development under different instructional conditions. In addition, it is important to determine what learners' internal differences as expressed in the verbal protocols may also reveal for both DOP and also different types of instruction, which could inform the L2 research community of which instructional type promotes a deeper or a more shallow processing. These gaps in the literature led the author to formulate the following research questions:

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

1. Do any of the three types of instructional conditions elicit deeper levels of processing as measured by think-aloud protocols? Hypothesis 1: Based on the absence of previous studies addressing DOP under different instructional conditions, the null hypothesis is adopted, that is, there will be no differences in terms of DOP among instructional conditions.
2. Is there a relationship between DOP, as measured by form–meaning connections in think-aloud protocols during a reading task, and subsequent interpretation and written production of the target form embedded in the reading task? Hypothesis 2: Based on the absence of previous published studies addressing the relation between DOP and subsequent interpretation and production of the target form, the null hypothesis is adopted, that is, there will be no relation between them.
3. Is there a relationship between the overall comprehension of a passage and DOP of the target form? Hypothesis 3: Based on previous findings by Leow et al. (2008) and Morgan-Short, Heil, et al. (2012), it is hypothesized that there will be a relationship between reading comprehension and DOP of the target structure.

METHOD

Participants

Participants were recruited from 12 sections of Spanish Intermediate classes at a medium-size research university on the East Coast. An initial pool of 140 participants signed up to participate in the study. In order to be included in the final sample and in the subsequent statistical analyses, (a) participants had to complete all three sessions of the study, (b) they could not have received formal instruction in a Romance language other than Spanish for more than 2 years, or (c) they could not have been previously exposed to the target form of the study. Fifty-two participants failed to fulfill the previous criteria, and a final sample of 88 participants ($n = 88$, mean age = 19.2) were included in subsequent analyses and were then assigned to the three different experimental conditions created for the current experiment.

Target form

The target form addressed in this study was the Spanish imperfect subjunctive in contrary-to-fact conditional clauses. An example of a conditional sentence and of the imperfect subjunctive (in bold) would be:

(1) *Si* ***aprobaras*** *los exámenes, estarías muy contento* (If you passed your exams, you would be very happy).

The imperfect subjunctive form *aprobaras* appears in the conditional clause introduced by *si*. This clause, as seen in (1), can precede the main clause *estarías muy contento* or it can also follow it reversing the order of the sentence, as in (2):

(2) *Estarías muy contento si aprobaras los exámenes* (You would be very happy if you passed your exams).

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

The imperfect subjunctive in contrary-to-fact conditional clauses was chosen as the target form because it is considered a challenging structure to be acquired by L2 learners. Numerous studies (Collentine, 1995, 1998; Fernández, 2008; Rosa & Leow, 2004a, 2004b; Terrell, Baycroft, & Perrone, 1987) have reported that at the intermediate level of Spanish, students still struggled to interpret and produce the present subjunctive in nominal clauses (e.g., Quiero que *vayas* a casa [I want you to go home/I want that you go home]). Collentine (1995) mentioned that one possible reason to explain this finding was that learners were not developmentally ready to learn the subjunctive in the third or fourth semester of Spanish. According to Collentine (1998), in a typical subjunctive instructional sequence, students see the subjunctive in nominal clauses first, followed by the subjunctive in adjectival clauses (e.g., Busco a un profesor que *hable* español [I look for a professor that speaks Spanish]), in adverbial clauses (Te llamaré cuando *llegue* a la Universidad [I will call you when I get to campus]), and in conditional clauses (see Examples (1) and (2) above). As the level of complexity increases, the structure where the subjunctive is embedded is seen at later learning stages. Timberlake (2007) posited that the acquisition of modality in Spanish is one of the most challenging aspects for L1 English speakers, because English lacks a specific form to refer to the subjunctive mood, employing in most cases the indicative present or preterite tenses (see subjunctive examples above). Other studies examining the acquisition of the Spanish subjunctive by native speakers of Spanish have found that this particular form is one of the last subjunctive forms to be acquired by children in a spontaneous manner, by the age of 7, after other forms such as the present subjunctive (Fernández & Aguado, 2007; López-Ornat, Fernández, Gallo, & Mariscal, 1994).

Materials

Instructions varying in their degree of explicitness prior to task. Three different types of instructional conditions varying in their degree of explicitness were employed in the current study:

1. The more explicit instructions (Appendix A) presented a detailed explanation defining the target form, the imperfect subjunctive, and how it is used within a conditional sentence. It also showed the endings for the different persons for the imperfect subjunctive for –AR verbs, presenting the whole conjugation of the verb *HABLAR*. It finally instructed that the conditional clause where the imperfect subjunctive is embedded can appear in first or second place in the sentence. The content was partially adapted from a grammar manual, *Repase y Escriba* (Dominicis & Reynolds, 2011), widely used in departments of Spanish at US universities.
2. The less explicit instructions (Appendix B) just let participants know that they would find a new grammar form in the reading passage that they had not seen before, and that they needed to try to figure out its form and use. They were presented a hint that drew their attention to the conditional *si* clause where the target structure was present.
3. The baseline instructions (Appendix C) just informed participants that they were going to read a passage without making any reference to the new target form.

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Experimental task

The experimental task (Appendix D) consisted of a modified version of an authentic reading passage. The length of the passage was of 415 words, a length in line with previous studies that employed a reading task to draw learners' attention to both meaning and form (Greenslade et al., 1999; Leow et al., 2008; Morgan-Short, Heil, et al., 2012). The target form, the imperfect subjunctive in conditional clauses, appeared 12 times throughout the text embedded in 12 different sentences. In 6 of these times, the conditional clause appeared first, and in the other 6, the order was reversed with the main clause appearing in the first place. The title of the passage was *Estudiar mejor: claves del éxito* (*Studying better: keys to succeed*), and in a nutshell, it offered students advice to improve their studying habits in order to obtain better results at school. An example of a piece of advice was: *Si cuidaras tu salud, te concentrarías mejor en los estudios*. This topic was expected to be familiar and of interest to university students. A first version of this passage was tested in a pilot test with students from a similar level of proficiency to account for any unknown vocabulary and overall difficulty of the test. Some slight modifications (sentence rewording and word modification) were made in order to better adjust it to participants' level of proficiency.

Learning assessment instruments

In this study learning was operationalized as the ability to interpret and produce the target form. For this matter, three different versions of interpretation (multiple choice; see Appendix E for an excerpt) and controlled production (fill-in the blank; see Appendix F for an excerpt) tests were administered to participants 2 weeks before treatment, immediately following treatment, and 2 weeks after treatment. Each test consisted of 30 items. Ten of these items targeted the target form and the target verbs (old items), 10 targeted the target form and alternate verbs (new items), and the remaining 10 were distractors, which targeted other grammatical aspects of Spanish (ser/estar and preterite/imperfect contrast, and the use of prepositions). Old items, new items, and distractors were randomized and presented in a different order for each version of the test.

Comprehension test

Because the experimental task asked participants to read for both meaning and form, a comprehension test (Appendix G) to account for understanding of the reading passage was included in this study. It consisted of 10 multiple-choice questions with three possible options per question. Upon observing previous results of a pilot test, 3 of the questions were slightly modified because they were deemed to be ambiguous and unclear for participants.

Think-aloud instructions

As explained more in detail in the Procedure section below, some participants in the present study had to think out loud while concurrently carrying out the

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Table 1. *Number of participants per condition before merging TA and NTA conditions*

	META	LETA	BLTA	MENTA	LENTA	BLNTA
Participants	15	15	15	13	16	14

Note: TA, Think aloud; NTA, non think aloud; ME, more explicit; LE, less explicit; BL, baseline.

Table 2. *Number of participants per condition after merging TA and NTA conditions*

	More Explicit	Less Explicit	Baseline
Participants	28	31	29

Note: TA, Think aloud; NTA, non think aloud.

experimental task. To that effect, a set of audiovisual think-aloud instructions on a Powerpoint was created for participants to practice thinking aloud and feeling more comfortable with the procedure before carrying it out during the actual task (see Appendix H for an excerpt of these instructions). The instructions consisted of a screen and audio recording of a Powerpoint where the instructions were also presented in a written mode, and where instructions were read aloud so that participants could both read and listen to them at the same time. Participants were first told that they had to think aloud all the time without trying to explain what they were thinking, but should just say it aloud clearly and loudly on the microphone. Then, they were presented with a fragment of a text, and one of the authors *modeled* a participant thinking aloud, saying out loud any thought that came to his mind while he was reading the text. In this modeling, the importance of thinking aloud, and not just reading the text aloud,[2] was emphasized. These instructions were recorded using QuickTime Player, and the total length was of approximately 4 min. One of the advantages of recording both the screen and the audio was that participants were forced to read and to listen to all of the instructions, not allowing them to skip any part of them.

Design

The study consisted of three different experimental conditions (Table 1),[3] varying in the degree of instructional explicitness received: more explicit, less explicit, and baseline (Tables 2 and 3).[4]

The way that the interpretation and controlled production tests were administered followed a randomized split-block design. Participants who were randomly assigned to complete Version A in the pretest, then completed Version B at posttest, and Version C at the delayed test. The other two possible sequences were B, C, A, and C, A, B. Approximately one-third of participants followed each of the sequences.

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Table 3. *Procedure*

Week 1 Session 1	Week 3 Session 2	Week 5 Session 3
~45 min	~60 min	~40 min
Controlled Produc. Pretests	TA Practice for TA Groups	Controlled Produc. Delayed Posttests
18 min max.	~4 min	18 min max.
Interpretation Pretests	Instructions	Interpretation Delayed Posttests
18 min max.	7 min max.	18 min max.
Learning Style Survey	Experimental Task (+TA) Reading Passage	Background Questionnaire
~15 min	9.25 min	~10 min
	Comprehension Test	
	10 min max.	
	Controlled Production Posttest	
	18 min max.	
	Interpretation Posttest	
	18 min max.	

Note: TA, Think aloud.

Procedure

The study materials and procedures were approved by the university internal review board (Study Number 2013-1064). All materials were presented on Blackboard. The advantages of using Blackboard to conduct the study were numerous. First, it allowed the researcher to assign a specific amount of time to the task in question and to choose which tasks were made available to participants for each session. Second, it also controlled for any potential *backtracking*, a limitation noted in Leow et al. (2008) and Morgan-Short, Heil, et al. (2012), keeping participants from going back to previous questions or to revisit the reading passage after submission. Third, another advantage was that all grades were automatically

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

saved into a gradebook and could be easily exported into an Excel sheet. Fourth and finally, the participants were very familiar with this software, because they use it regularly for their classes.

The study consisted of three sessions (see Table 3 for procedure and times[5]). All sessions took place in a language laboratory equipped with Mac computers. Session 1 took place 2 weeks before treatment. Participants came to the laboratory and completed, in this order, the informed consent to participate voluntarily in the study, the controlled-production, and interpretation pretests.[6]

At treatment, 2 weeks after Session 1, participants were randomly assigned to the experimental conditions and completed the tasks in the following order: (a) ±think-aloud practice/instructions, (b) instructions varying in explicitness, (c) experimental task (reading passage + think-aloud), (d) comprehension test, and (e) production and interpretation posttests. Those assigned to the silent conditions followed the same order of tasks starting at (b). At the end of Session 2, participants signed up for Session 3, but were not told what it would be about.

In Session 3, 2 weeks after treatment, participants completed the controlled-production and interpretation delayed posttests and then the background questionnaire.

Scoring and coding

For comprehension, interpretation, and controlled production tests, participants were given 1 point for each right answer provided in the tests and 0 for a wrong answer. Responses had been assigned on Blackboard. For the controlled production tests, in addition to a strict scoring, a lenient scoring[7] was considered when participants provided the target form in one of the two following ways: (a) right tense but wrong person agreement (e.g., tú *hablara* instead of tú *hablaras*); or (b) right tense with a spelling error (e.g., él *entenara* instead of él *entrenara*). Thus, participants showed they had acquired a partial understanding of the target form, even if they were not able to produce the target form in a strictly correct manner.

Therefore, two separate scoring scales (strict and lenient) for the controlled-production tests were computed for the analyses conducted later on in the study. Participants' instances of processing found in the think-aloud data were categorized into three different levels: low, intermediate, and deep, assigning them a 1, 2, or 3 value, respectively (e.g., Leow et al., 2008; Morgan-Short, Heil, et al., 2012; Hsieh et al., 2015). Criteria to decide which level they would fall into were slightly adapted from Leow et al. (2008) and Morgan-short, Heil, et al. (2012; see Table 4) to fit with the type of experimental task employed in this study. Considering the qualitative nature of the data, 20% of the think-aloud transcriptions were coded and compared with two additional raters to calculate the interrater reliability. These two raters were graduate students in applied linguistics at the university where the data were collected by the researcher, and they already had some familiarity with the collection and coding of qualitative data. They were explained how to code the transcriptions following the criteria detailed in Table 4. They were coded separately by the two raters and by the researcher, and then the three got together to compare the resulting individual codings. Prior to that, the researcher and the two additional raters had listened to three think-aloud files (one

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Table 4. *Operationalization of depth of processing (DOP) grammatical items*

	Low DOP	Intermediate DOP	Deep DOP
Description	Shows no potential for processing target form grammatically	Comments on target item in relation to grammatical features	Arrives at a partially, fully, or (in)accurate target underlying grammatical rule
Descriptors	• Reads target quickly • Translates the phrase to English but leaves the target in Spanish • Carefully pronounces target item • Repeats target item • Significant pause before or after the target form • Says s/he is not sure what it is • Does not spend much time processing target item • Low level of cognitive effort to process target item grammatically	• Spends a bit more time processing target item • Makes comments that indicate some processing of target item • Carefully emphasizes target word or structure (including rising intonation or prolonging some sounds or syllables, or saying "oh!") • Arrives at a partial translation with some cognitive effort • Some level of cognitive effort to process target item grammatically	• Makes hypotheses regarding target item • Provides a partially, fully, or (in)accurate target underlying grammatical rule • Provides correct translation of target form • Corrects, or ostensibly attempts to correct, a previous translation • Spends much time processing target item • High level of cognitive effort to process target item (un)grammatically

Table 5. *Individual raters' agreement and total agreement*

	R1–R2	R1–R3	R2–R3	Total
Rating agreement	$r = .100$	$r = .100$	$r = .927$	$r = .976$

per condition) to revisit the descriptors and better adapt them to the content of the think-aloud transcriptions. The final resulting agreement was high: 97.57% (see Table 5 for individual raters' agreement).

RESULTS

The results for the research questions (RQs) of the study are reported in this section. Statistical analyses for RQ1 and RQ3 were performed using SPSS 22. Analyses for RQ2 were performed using STATA 13. The α level was set at 0.05 throughout all analyses.

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Table 6. *Number of instances of processing per level and per condition*

	Condition			
Processing	More Explicit	Less Explicit	Baseline	Total
Low	28	33	58	119
Intermediate	114	121	102	337
Deep	31	12	5	48
Total	173	166	165	504

Note: Participants: more explicit = 15, less explicit = 14, baseline = 14.

Table 7. *Descriptive statistics of Processing instances per level and per Condition*

	Condition		
	More Explicit	Less Explicit	Baseline
Processing	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Low	1.87 (2.32)	2.36 (1.78)	4.14 (2.8)
Intermediate	7.6 (3.02)	8.64 (1.86)	7.28 (2.67
Deep	2.07 (2.68)	0.86 (1.46)	0.21 (0.58)

Note: Participants: more explicit = 15, less explicit = 14, baseline = 14.

Preliminary analyses

Before addressing the study RQs, preliminary analyses were run to ensure that there was homogeneity at pretests among groups. In other words, the goal of these analyses was to show that any existing differences observed after treatment were a result of the treatment itself, and not of inherent differences among conditions. A one-way analysis of variance was performed with the interpretation pretest scores as dependent variable and the experimental conditions as fixed factor. The analysis revealed no significant differences among groups, $F(5, 71) = 0.656$, $p = .66$, partial $\eta^2 = 0.044$, observed power = 0.23. No statistical analysis was run for the controlled production pretest, because a visual inspection observed that most participants across conditions scored 0, and an overall homogeneity was assured. These results granted a significant homogeneity among groups prior to treatment at both interpretation and production pretest scores.

Response to RQ1: Do any of the three types of instructional conditions elicit deeper levels of processing as measured by think-aloud protocols?

In order to respond to RQ1, a nonparametric test, a 3 × 3 chi-square (condition × level of processing), was performed to address the frequency of instances of processing per level and per condition (see Tables 6 and 7 for a detailed count

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

and for the descriptive statistics respectively), which rendered a significant result,[8] χ^2 (4) = 36.907, p = .00, with a medium effect size, v = 0.27 (Larson-Hall, 2010). Forty-three think-aloud protocols (more explicit: 15, less explicit: 14, baseline: 14) produced a total of 504 instances of processing (examples for each of the levels of processing can be seen in Appendix I). As seen in Table 6, the more explicit condition produced a larger number of instances of deep processing, followed by the less explicit and by the baseline conditions. When instances of intermediate processing were compared, the three conditions produced an approximately similar number of instances, with the less explicit condition producing slightly more than the more explicit and the baseline groups. With regard to the instances of low processing, the baseline condition produced more instances than the less explicit and the more explicit conditions. If the total number of instances of processing, regardless of its depth, is considered, all three conditions appeared to produce a roughly equal amount.

From these numbers, it seems that a more explicit instructional type was related to more instances of deep processing, and as the instructional explicitness decreases, so does DOP, yielding more instances of lower processing.

Response to RQ2: Is there a relationship between DOP, as measured by form–meaning connections in think-aloud protocols during a reading task, and subsequent interpretation and written production of the target form embedded in the reading task?

In order to respond to RQ2, a generalized linear mixed model analysis was performed using Stata, version 13 (see http://www.stata.com/stata13/).[9] This analysis was conducted with Stata due to the nonindependent nature of the data and to the binary coding (1 or 0) of the dependent variables. Prior to running this analysis, an item-by-item (IBI) codification had been conducted for each response provided for old items at the interpretation and controlled production posttests. These were coded as either interpreted or not interpreted, and assigned a value of 1 or 0 (variable name: interpretation IBI); produced or not produced (strict grading), and assigned a value of 1 or 0 (variable name: strict production IBI); and produced or not produced (lenient grading), and assigned a value of 1 or 0 (variable name: lenient production IBI). The think-aloud protocols were then examined to observe which level of processing (low, intermediate, or deep) participants had reached for each particular instance (variable name: DOP IBI). There were 421 instances of processing out of a possible total of 430 (43 participants × 10 items). Nine of the instances were not *verbalized* by participants. Because no evidence of processing (at any level) could be extracted from the think-aloud data for those items, they were not included in the analysis. The analysis (see Table 8) yielded a nonsignificant correlation between DOP IBI and Interpretation IBI (b = –0.58, p = .56, confidence interval = 0.48–1.49); a significant correlation between DOP IBI and strict production IBI (b = 2.47, p = .01, confidence interval = 1.20–4.82), and a nonsignificant correlation between DOP IBI and lenient production IBI (b = 1.59, p = .11, confidence interval = 0.86–4.16). These findings would indicate that as DOP increased, it was more likely that participants produced the target form right. However, a deeper processing was not significantly related to a

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Table 8. *Correlations between depth of processing and interpretation and production of the target form*

IBI	Coefficient *b*	*SE*	*p*	Confidence Interval
Interpretation	–0.58	0.24	.56	0.48–1.49
Strict production	2.47	0.85	.01*	1.20–4.82
Lenient production	1.59	0.76	.11	0.86–4.16

Note: Participants = 43. IBI, Item by item.
*$p < .05$.

correct interpretation of the target form or to an accurate production when graded leniently.

Response to RQ3: Is there a relationship between the overall comprehension of a passage and DOP of the target form?

In order to respond to RQ3, a two-tailed Pearson correlation was run between participants' level of processing and comprehension scores. In order to identify each participant's level of processing, an examination of all instances of processing provided by the participant was conducted, and the most frequently reported level was selected. Overall, 6 participants (14% of the total) were coded for Level 1 (low), 34 (79%) were coded for Level 2 (intermediate), and 3 (7%) were coded for Level 3 (deep). The correlation analysis ($r = .37$, $p = .013$) rendered a significant positive relation between the variables: the deeper the processing as revealed by the think-aloud data, the higher the score obtained at the comprehension test.

DISCUSSION

Discussion for RQ1: Do any of the three types of instructional conditions elicit deeper levels of processing as measured by think-aloud protocols?

Upon evaluation of the results, the null hypothesis was rejected, as the more explicit condition yielded more instances of deep processing than the other two conditions. As the instructional explicitness decreased, so did the level of processing of the instances of the target forms, as indicated in the think-alouds. In other words, participants who received metalinguistic information about the target form were able to provide more instances of deeper processing, such as (in)accurate metalinguistic comments, or correct translations to English of the target form and the conditional structure.

Previous studies addressing the specific relationship between DOP and different degrees of instructional explicitness in the L2 have not been identified. Within the construct of awareness, some studies have examined this relation (Rosa & Leow, 2004a; Rosa & O'Neill, 1999). Rosa and O'Neill (1999) found that participants who received explicit information about the target form produced more instances

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

of awareness at the level of understanding than participants in the less explicit condition, but the difference was not significantly different. This difference in instances at the level of understanding did not relate to subsequent performance, as the more explicit condition did not outperform the rule-search condition. Rosa and Leow (2004a) found that participants in their more explicit conditions did report higher levels of awareness, and these levels of awareness were related to higher accuracy at intake and knowledge after task, as the more explicit conditions outperformed the less explicit ones. As the authors put it: "When only one source of information was available to learners, techniques such as a pre-task or explicit feedback were more effective than implicit feedback in helping learners develop higher and more sophisticated levels of awareness" (p. 286).

In addition to the awareness at the level of understanding, they also found that awareness at the level of noticing still promoted learning, supporting Leow (1997), but contradicting Robinson (1996, 1997). Robinson found that only awareness at the level of understanding promoted learning. Rosa and Leow (2004a) argued that the lack of task essentialness in Robinson's studies and the different assessment tools used (a grammaticality judgment test instead of an interpretation test) could explain that participants in his studies needed to reach a higher level of awareness to show an impact on learning.

Within the limitations of comparing levels of awareness and levels of processing, current findings would be in line with Rosa and Leow (2004a) and Rosa and O'Neill (1999) in that the more explicit condition produced a higher number of instances of deep processing (19 more than the less explicit condition). However, if we examine these findings in light of the performance by condition immediately or 2 weeks after treatment, this deeper processing extracted from the think-aloud data for the more explicit condition did not translate into an equal L2 development when compared to the less explicit group. The less explicit group outperformed the more explicit one, especially at the delayed posttest. It could be argued that an intermediate processing of the imperfect subjunctive, where instances were roughly similar between the two conditions, is sufficient to trigger learning of the target form. The long-term effects of learning under explicit or more implicit conditions would require an alternative explanation. Research seems to suggest that knowledge derived from implicit, or less explicit, learning conditions produces longer term effects. In contrast, explicit learning condition effects will hold for the short term, but will show a decay in the long term. Participants receiving pretask grammar information produced the accurate mood and tense equally or slightly better than those not receiving it right after treatment. However, that explicit knowledge of the imperfect subjunctive was not maintained 2 weeks after treatment. Further evidence was provided by Morgan-Short, Finger, Grey, and Ullman (2012), which yielded neurolinguistic evidence of a more nativelike processing for their implicit condition in the long term. Whereas this neural evidence cannot be necessarily compared to an intermediate or deep level of processing, it would probably indicate a cognitive effort mechanism more in line with native patterns of processing. If we consider the baseline condition significant improvement in productive abilities from pretest to the delayed test and we take a look at their instances of intermediate processing, 102, not that distant from the 121 (less explicit) and 114 (more explicit), it could further suggest that processing

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

the target form deeply is not strictly necessary to promote L2 development of the imperfect subjunctive in a conditional clause.

Discussion for RQ2: Is there a relationship between DOP, as measured by form–meaning connections in think-aloud protocols during a reading task, and subsequent interpretation and written production of the target form embedded in the reading task?

The answer for this question would be negative with respect to the interpretation of the target form, and positive with respect to the production of the target form. In other words, the deeper the processing of a specific target form, the more likely it was for that participant to have gained knowledge of both the subjunctive mood and the imperfect tense. On the contrary, this same likelihood was not observed when measuring the intake of the target form. Therefore, the null hypothesis was rejected.

Schmidt's (1993) model of attention in SLA proposed the *noticing hypothesis*, by which he posited that attention controls access to awareness and is responsible for *noticing*, which is the necessary condition for the conversion of new input into intake. Robinson (1995) also suggested that awareness is a key element in the subsequent intake and knowledge of a lexical or a grammatical item. Therefore, the input cannot be processed at a minimum level to be fully internalized, but will require a higher cognitive effort by the learner.

Within the construct of awareness, empirical studies have provided evidence to support Schmidt's hypothesis (Leow, 1997, 2001; Rosa & Leow, 2004a; Rosa & O'Neill, 1999). Leow (1997, 2001), at the level of item learning, investigated how different levels of awareness related to the recognition and controlled production of several morphological forms on immediate posttests. Participants who reported higher levels of awareness performed significantly better at gains of intake and knowledge of the new structure. Rosa and O'Neill (1999), only at the level of intake, and Rosa and Leow (2004a), for both intake and knowledge, yielded similar trends. Martínez (2010) also found a correlation between levels of awareness and recognition and production of a grammatical target form for both old and new items. The form targeted was the present subjunctive used in adverbial clauses introduced by *cuando* (when), which, as the target form in the present study, makes reference to an action happening in a hypothetical future.

Within the construct of DOP, Calderón (2014) recently found a positive relation among a deeper processing of the grammatical target form, the preterite third-person singular, and subsequent knowledge of the form in a controlled production test. However, she found no relationship with the subsequent intake of the form.

The present findings would be in line with Calderón's (2014) and would partially support Leow (1997, 2001), Rosa and Leow (2004a), and Martínez (2010). In order to explain why this relation has only been found with learners' productive abilities, we could argue that a deeper level of cognitive effort is necessary to promote successful knowledge of the target form, while in order to intake the latter, a lower level appears to be sufficient. Considering the L2 learning process, where intake of a new form/structure comes at an earlier stage than knowledge of that form, a deeper processing seems to be necessary to transform that intake

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

into knowledge within this learning system. However, this assumption does not seem to be correlated with the complexity of the target form, as Calderón's target form (simple) would differ from Martínez and the present study's target forms (complex). When referring to learners' productive abilities, it is also revealing to examine the contrast between the accurate production of both mood and tense or just mood. That the correlation between deeper levels of processing and the production of the subjunctive mood (but not of the right tense) did not turn out to be significant could indicate that a lower level of processing is sufficient to produce a quasi-accurate target form (e.g., tú *hablara* instead of tú *hablaras*), but a deeper processing will be necessary to produce a total accurate version that combines the right mood and tense. In sum, present findings seem to be in line with previous research that has addressed the relation between higher and deeper levels of awareness and processing, and subsequent measurement of intake and knowledge of the target form.

Discussion for RQ3: Is there a relationship between the overall comprehension of a passage and DOP of the target form?

The response for RQ3 would be positive, and the hypothesis confirmed. Deeper levels of processing were correlated with a better comprehension of the reading passage. These findings would be in line with Morgan-Short, Heil, et al. (2012)[10] and Calderón (2014), who found that participants performing a dual task (reading a passage for form and meaning), showed higher comprehension scores when their overall processing had been deeper. Leow et al. (2008), while not making any strong conclusion due to their small sample size, had suggested that the marginal differences found among experimental conditions in their study could be due to the low number of instances of processing reported.

In the field of cognitive psychology, Craik and Lockhart's (1972) seminal study on levels of processing, revisited by Craik (2002), suggested that deeper levels of processing were necessary for a better "recall" of information. In other words, a deeper processing would entail a better recognition of meaning, while a "shallow" processing, that of surface aspects, would not be sufficiently deep for the participant to recall the meaning he had been exposed to. In SLA, Norris and Ortega (2000) posited that instructional conditions seemed to be more effective when they addressed learners' attention to both meaning and form, rather than just meaning or form alone.

However, these findings would refute VanPatten's (2004) *primacy of meaning* principle, which stated that L2 learners, when exposed to input, will process meaning rather than form, and that a dual task where attention has to be allocated to both of them will affect negatively the comprehension of the input. It is important to note that the present study, as well as Morgan-Short, Heil, et al. (2012), and Calderón (2014), provided written input to participants. However, VanPatten (1990) found evidence that supported his principle when input was presented in the aural modality. In that study, attention to both meaning and form decreased participants' comprehension of the listening passage. Two other studies that replicated VanPatten's yielded different patterns with respect to the dual task impact on comprehension: Greenslade et al. (1999), in the written modality, found

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

evidence to support VanPatten; Wong (2001), who presented the input in both the aural and written modalities, supported VanPatten's results when attending to the aural modality, but rendered support for findings from the present study when participants received written input (cf. Leow et al., 2008).

Révész and Brunfaut (2013) also addressed a receptive skill, learners' listening abilities, and how task characteristics could affect task difficulty. One of these characteristics was syntactic complexity, whose subcategories included subordination and phrasal complexity. Neither of these components was found to predict task difficulty. Keeping in mind the differences between listening and reading modalities, we could argue that syntactic complexity could be the main difficulty of the current experimental task. Participants were exposed to a number of subordinate sentences containing an unknown grammar structure and form. Révész and Brunfaut added that learners identified syntactic complexity as a clear predictor of task difficulty, which did not correspond with their performance. Based on this premise, it could be hypothesized that learners would also consider the syntactic complexity of the reading passage with the new target forms embedded in conditional sentences as a main indicator of its task difficulty, and that could be detrimental for their comprehension of the passage, in line with VanPatten's argument. However, this was not the case, and Révész and Brunfaut suggested that participants might have actually allocated a higher cognitive effort to comprehending the meaning of the text in addition to processing and gaining knowledge of the new target form, as they were aware in two of the three instructional conditions of the presence of new grammatical information in the reading passage.

In summary, it appears that the type of presentation modality of the input (written vs. aural) could play a role in decreasing comprehension, as seen above. However, except for Leow et al. (2008), Morgan-Short, Heil, et al. (2012), and Calderón (2014), no claims or comparisons can be made with the rest of studies with regard to how DOP relates to comprehension scores, as they did not address the variable of DOP in their designs.

Limitations

As with any empirical study, some limitations should be noted. First, the lack of reactivity, as it was controlled in the present design, could be due to the highly controlled characteristics of the testing instruments employed, as previous reactivity has been found, for the most part, in less controlled materials (Bowles, 2010). Different testing tools that measure participants' *free* production, for instance, could be more likely affected positively or negatively by the use of think-aloud protocols. An additional limitation related to verbal protocols is the issue of *veridicality*, or lack of, referring to the omission of information by participants thinking aloud. It cannot be taken for granted that they verbalized every single line of thought in their minds, and some instances of processing, lower or deeper, might have been missed.

Second, another limitation is related to the effect that repeated exposure (three times) to testing materials can have on the learning process, especially for participants at the baseline condition. Although the time between sessions (2 weeks) and the three different versions of the tests would reduce that exposure effect, it

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

cannot be completely ruled out that this exposure added to the gains obtained from exposure to the reading passage at the treatment stage.

Conclusion

The present study intended to further explore the line of research initiated by Leow et al. (2008) and continued by Morgan-Short, Heil, et al. (2012), addressing whether instructional conditions differing in explicitness have an effect on DOP found as participants are exposed to a complex grammar form, the imperfect subjunctive, in a reading passage. Overall, results on processing yielded instances of deeper processing for the more explicit condition, and this processing became more shallow as the instructional explicitness decreased. When compared to the performance yielded by each condition, it could be suggested that reaching an intermediate level of processing might be sufficient to trigger the learning of the target form. In addition, deeper processing was positively correlated with a subsequent accurate performance at production (but not interpretation), and with higher scores at comprehension. The latter findings would support previous studies that also found a positive correlation between deeper levels of processing and posttask performance at comprehension (e.g., Calderón, 2014 Morgan-Short, Heil, et al., 2012) and at controlled production (Calderón, 2014).

APPENDIX A

Instructions for participants in the more explicit condition

You are going to read a passage in Spanish, and then you will be asked some questions in English to see how well you understood the passage. You will be asked to think aloud while you read the passage (as you did in the practice exercise). Also, there is a new structure in the passage that you haven't learned in class yet. So I will explain it to you.

The new structure is called "past subjunctive," and it is used in hypothetical (contrary-to-fact) situations. It is formed by taking the stem from the third-person plural form of the preterite minus *–on.*

So if the third-person plural form of the preterite of *hablar* is *hablaron*, we would eliminate *-on* and have *hablar-*. We would add the past-subjunctive endings **–a**, **–as**, **–a**, **–amos**, **–ais**, **–an**, and we would have the following forms:

hablara habláramos
hablaras hablarais
hablara hablaran

When we have a situation that is perceived as hypothetical (contrary to fact), the past subjunctive is used in the *if* clause (or hypothetical clause) and the conditional is used in the result clause (or main clause). The clauses can occur in either order. Observe the following sentences:

Si **tú participaras** más en clase, **recibirías** una nota más alta.
Juan se **graduaría** en mayo si **aprobara** la clase de español.

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Ok, now we are ready. You are going to read the passage in Spanish where this new structure I have just explained to you will appear several times, and then you will answer some comprehension questions in English based on the text. And remember you have to think aloud while you read the passage!

Good luck!

APPENDIX B

Instructions for participants in the less explicit condition

You are going to read a passage in Spanish, and then you will be asked some questions in English to see how well you understood the passage. You will be asked to think aloud while you read the passage. Also, there is a new structure in the passage that you haven't learned in class yet. So your goal will be to try to figure out the rule.

You will see it several times throughout the passage. Your goal is to figure out the grammar rule or rules for that structure. Paying attention to the sentences that contain a *si* clause might help you!

Ok, now we are ready. You are going to read the passage in Spanish, and then you will answer some comprehension questions in English based on the text. And remember you have to think aloud while you read the passage!

Good luck!

APPENDIX C

Instructions for control participants

You are going to read a passage in Spanish, and then you will be asked some questions in English to see how well you understood the passage. You will be asked to think aloud while you read the passage.

Ok, now we are ready. You are going to read the passage in Spanish, and then you will answer some comprehension questions in English based on the text. And remember you have to think aloud while you read the passage!

Good luck!

APPENDIX D

Estudiar mejor: Claves del éxito

Texto adaptado de: http://centros5.pntic.mec.es/ies.de.bullas/revista/estudiar.htm

La sabiduría popular dice que el éxito en la vida depende de dos ingredientes: inteligencia y ganas de triunfar, pero que el primero fracasa sin el segundo: lo importante es la voluntad. Los expertos en educación añaden que si no se motivara a los estudiantes lo suficiente, sólo su inteligencia no bastaría para tener éxito. Cualquier ser humano que sepa leer es capaz de aprender, pero aprendería más rápido si trabajara con más esfuerzo. Es cierto que algunas personas parecen haber nacido dotadas para el estudio, ya que asimilan y retienen los datos más rápidamente que otras, pero en general aprender es, simplemente, una cuestión de disciplina. Incluso un superdotado físico no conseguiría ni un triunfo deportivo si no

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

entrenara muchas horas cada día. Aprender es sencillo, sólo hay que proponérselo. Y para empezar, nada mejor que establecer cuáles son tus prioridades.

"Mens sana in corpore sano." Cerebro y cuerpo forman el mismo conjunto, y cuando estás enfermo o débil pierdes capacidad para pensar. Si cuidaras tu salud, sería más fácil concentrarte en los estudios. Del mismo modo, serías más productivo si practicaras ejercicio físico de manera regular. De igual manera, si descansaras una media de 8 horas, tu cerebro se fortalecería y procesaría mejor la información.

El primer paso al disponerte a estudiar es motivarte, poner toda tu energía en el empeño de aprender. Estructurarías mejor tu aprendizaje si te marcaras unos objetivos a corto plazo. Ten un buen nivel de autoestima: incluso si alguna asignatura te causara dificultades, deberías confiar en ti. Fortalece tu sentido de superación: si te preocupara no mejorar con rapidez, deberías pensar en las ventajas de tu esfuerzo. Cada día, al ponerte a estudiar, repasa lo aprendido el día anterior, y comprobarás lo que realmente has avanzado. Concéntrate poniendo los cinco sentidos en el estudio. Si no controlaras tus emociones, podrían restarte energía y concentración, como la tensión, la angustia y la ansiedad. Supéralas relajando tu cuerpo.

Es importante estudiar siempre en el mismo sitio porque la familiaridad con el entorno ayuda. Crea tu propio espacio cerca de una ventana, instala una mesa grande para distribuir apuntes, libros. Si estudiaras cada día en un sitio diferente, te distraerías con mucha facilidad. Es importante el orden en tu mesa para que no pierdas el tiempo. Finalmente, el frío impide concentrarse y el calor produce agotamiento, por lo que lo ideal es conseguir unos 64 o 68 grados. Te costaría mucho concentrarte si la temperatura cambiara con mucha frecuencia.

APPENDIX E

Interpretation test (excerpt)

You are going to read some situations and then you will need to select the option that better describes it.

1. Pablo is a senior in high school, and he is not sure he wants to attend college. He should consider doing something different if he is not very motivated.
   a. Si a Pablo no le motivara asistir a la universidad, debería hacer algo diferente.
   b. Si a Pablo no le motivaría asistir a la universidad, debería hacer algo diferente.
   c. Si a Pablo no le motivaba asistir a la universidad, debería hacer algo diferente.
2. Juan is a disciplined student who goes to class every day. However, he is not that disciplined when it comes to doing his homework, and his results are not that good.
   a. Juan conseguiría mejores resultados si trabajaba con más regularidad.
   b. Juan conseguiría mejores resultados si trabajara con más regularidad.
   c. Juan conseguiría mejores resultados si trabajará con más regularidad.
3. María is a physically gifted athlete at Georgetown, and she trains with less intensity than her teammates. She knows she needs to train harder to win more races.
   a. Si María entrenaría más, ganaría más carreras.
   b. Si María entrenó más, ganaría más carreras.
   c. Si María entrenara más, ganaría más carreras.

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

4. You do not take very good care of your health, and get sick quite often. You are aware that taking better care of your health would boost your academic performance.
   a. Mejorarías tu rendimiento académico si cuidarás tu salud.
   b. Mejorarías tu rendimiento académico si cuidaras tu salud.
   c. Mejorarías tu rendimiento académico si cuidaste tu salud.
5. You used to do exercise every day. Nowadays, you barely do any exercise and feel very tired. You wish you could do more exercise to have more energy.
   a. Si practicaras ejercicio físico con frecuencia, serías más productiva.
   b. Si practicabas más ejercicio físico con frecuencia, serías más productiva.
   c. Si practicarás más ejercicio físico con frecuencia, serías más productiva.
6. You do not sleep more than 5–6 hours at night, and as a result, you are tired all day long. You know that being able to sleep more hours is good for your productivity.
   a. Procesarías mejor la información si lograbas dormir más horas.
   b. Procesarías mejor la información si lograrías dormir más horas.
   c. Procesarías mejor la información si lograras dormir más horas.
7. María has low self-esteem because she is struggling in some of her courses. A higher confidence in herself is very important for her to succeed.
   a. María debería tener más confianza si alguna asignatura le causaba dificultades.
   b. María debería tener más confianza si alguna asignatura le causará dificultades.
   c. María debería tener más confianza si alguna asignatura le causara dificultades.
8. Roberto works very hard on his courses but with little improvement, which worries him. He knows that being more positive will be better for him.
   a. Si a Roberto le preocuparía mejorar muy despacio, debería ser más positivo.
   b. Si a Roberto le preocupara mejorar muy despacio, debería ser más positivo.
   c. Si a Roberto le preocupó mejorar muy despacio, debería ser más positivo.
9. You normally get very tense and anxious before finals. Your advisor recommended you to do some relaxation exercises to increase your concentration.
   a. Aumentarías tu concentración si controlaras tus emociones.
   b. Aumentarías tu concentración si controlarás tus emociones.
   c. Aumentarías tu concentración si controlabas tus emociones.
10. You love studying at different places, at home, at the library, at the cafeteria, but you do not know it actually distracts you even more than studying at one place.
    a. Si estudiaras en el mismo lugar todos los días, no te distraerías tanto.
    b. Si estudiaste en el mismo lugar todos los días, no te distraerías tanto.
    c. Si estudiarías en el mismo lugar todos los días, no te distraerías tanto.

APPENDIX F

Controlled written production test (excerpt)

1. The Spanish instructor knows he needs to motivate Pablo more so that he is able to pass his exams in May.
   Si el profesor no (MOTIVAR) _____ a Pablo, éste no aprobaría los exámenes.
2. Luis has a lot of stress when he works at his office. He knows that teleworking from home could be better for him.
   Luis tendría menos estrés si (TRABAJAR) _____ desde casa.

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

3. Pablo is a great swimmer, but he knows he needs to keep training really hard to stay at the top.
   Si Pablo no (ENTRENAR) _____ más horas, no sería el mejor.
4. You have a very good friend at your hometown, but you have not really called her lately. You know you should call her more often to keep your friendship alive.
   Mantendrías una muy buena amiga si (CUIDAR) _____ tu amistad.
5. You have midterms in 2 weeks, and you are starting to feel stressed. You might sign up for yoga classes to feel more relaxed.
   Si (PRACTICAR) _____ yoga, te sentirías más relajado.
6. Your roommates are very noisy, and they do not let you sleep at night. You want to talk to them so that you can sleep more hours.
   Serías más productivo si (DESCANSAR) _____ más horas.
7. Juan is not good at math, and might struggle at his calculus class. He is going to try to remain positive and self-confident.
   Juan intentaría ser optimista si la clase de cálculo le (CAUSAR) _____ problemas.
8. Pablo has a lot of pressure from his parents to get straight A´s. He is afraid he might not be able to achieve that goal, but he is going to remain calm.
   Si a Pablo le (PREOCUPAR) _____ no obtener A´s en todas sus clases, debería permanecer tranquilo.
9. You are a very emotional student, and you know that your emotions make it difficult for you to concentrate, and that you should control them.
   Te concentrarías mejor si (CONTROLAR) _____ tus emociones.
10. You find it hard to study at the library, and you think you might focus better on your studies in your room.
    Si (ESTUDIAR) _____ en tu habitación, aprenderías más rápido.

APPENDIX G

Comprehension test

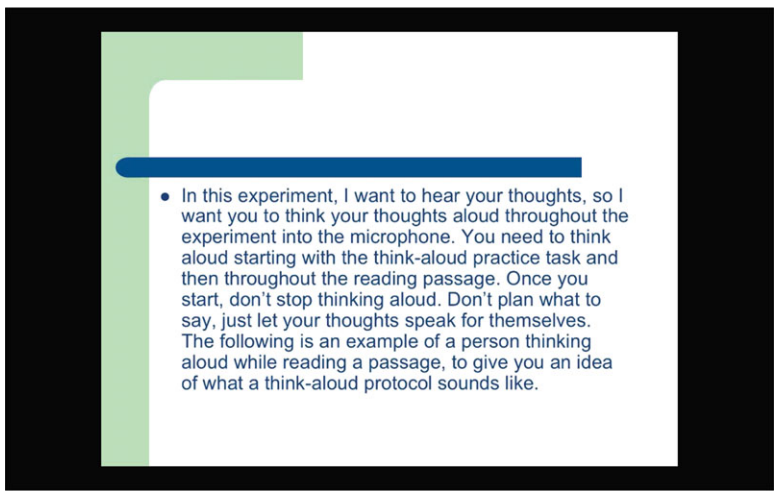
Based on what you have just read, circle the right answer.

1. According to popular knowledge . . .
   a. Intelligence is what really matters to succeed in life.
   b. Intelligence and willingness to succeed are equally important.
   c. Willingness to succeed is what really matters to succeed in life.
2. Being able to learn . . .
   a. occurs if we know how to read, despite an additional effort.
   b. occurs if we know how to read and put some additional effort.
   c. is not related to either (a) or (b).
3. A highly gifted individual . . .
   a. will not need as much discipline to succeed as an average individual.
   b. will need the same discipline to succeed as an average individual.
   c. will need more discipline to succeed than an average individual.

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

4. *Mens sana in corpore sano* refers to . . .
   a. the importance of doing exercise before studying to obtain better results.
   b. the fact that doing physical exercise is actually more important than studying.
   c. the importance of physical exercise as a complement to mental exercise.
5. Students would better process the information if . . .
   a. they took short breaks every 2–3 hours of study.
   b. they ate small snacks (like energy bars) that enhance productivity.
   c. they slept between 7 and 9 hours a day.
6. Better structuring the information will depend on . . .
   a. setting long-term goals.
   b. setting short-term goals.
   c. a combination of both.
7. In case learning is not happening at the desired pace . . .
   a. students will have to modify their studying habits.
   b. students will need to value their effort.
   c. students will need to seek external advice.
8. Not controlling their emotions . . .
   a. will negatively affect their energy and concentration.
   b. will affect their sleep and subsequent performance.
   c. will negatively affect their relations with other students.
9. Studying always at the same place . . .
   a. might become tedious, discouraging the student.
   b. could actually enhance productivity.
   c. might be distracting for the student, reducing productivity.
10. The temperature to study . . .
    a. should be a bit cool to help the student better concentrate.
    b. will have an impact on the student's performance.
    c. will not really have an impact on the student's performance.

APPENDIX H



https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

APPENDIX I

Instances of processing

Deep level of processing. "if you eat salad, if you think, no, its not salad geez, salad salud seria mas facil concentrarte en los estudios What?! If . . . so that´s the new thing, the hypothetical . . . if you seria mas facil . . . its easier to concentrate on your studies, so if you do something its easier. . . ." (Participant 16)

"if one is not motivated, or if one does not motivate students, if one were not to motivate students enough only their intelligence no bastaría, would not be enough to have success." (Participant 43)

Intermediate level of processing. "If you worry yourself you won't umm improve your speed, you should think in the . . . Ventajas de tu esfuerzo, I don't know what ventajas is, maybe windows? You'll see into something?" (Participant 49)

"If you don't control your emotions podrian . . . put . . . you should put . . . or put . . . no you should . . . no you can restart, regain energy and concentration." (Participant 52)

Low level of processing. "Include si alguna asignatura te causara dificultades, deberías confiar en ti." (Participant 37)

"and study every day in a different site, distra . . . I think you'll be distracted." (Participant 21)

ACKNOWLEDGMENTS

I thank Dr. Ronald P. Leow for his feedback on a previous version of this manuscript and Ariel Zach and Pablo Camus for their help with the transcription and coding of verbal protocols. Special thanks to the three anonymous reviewers for their insightful and constructive comments.

NOTES

1. Leow et al.'s model (2015) is the first one to postulate a key role for DOP at this stage.
2. A different set of instructions was tested in a pilot test. While transcribing the think-aloud files, it came to the researcher's attention that some participants had just read aloud the text. This was the main reason why this new set of instructions was designed and implemented in this study.
3. In the original study, there were initially six experimental conditions (three silent and three thinking aloud) in order to control for reactivity (Table 1). One-way analyses of variance were conducted to compare performance at comprehension, interpretation, and production, as well as time on task. Only a latency effect was found for time on task, with the think-aloud groups spending more time on task than the silent conditions, $F\ (5, 82) = 12.561$, $p = .00$, partial $\eta^2 = 0.434$, observed power $= 1.00$. However, as no reactivity emerged for performance at any type of test, the think-aloud and silent conditions were merged into the three conditions differing in explicitness reported in the present study (Bowles, 2010). The previous time on task differences

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

disappeared when these three conditions were compared, F (2, 85) = 0.95, p = .39, partial η^2 = 0.022, observed power = 0.21.

4. As this condition was exposed to meaningful input when reading the passage, it cannot be considered a *pure* control condition, but a baseline condition.
5. Times for instructions and tasks were piloted and slightly modified for this study.
6. As one of the reviewers rightly noted, it is important to inform readers about the experimental groups' performance at the interpretation and production post- and delayed posttests, even if it is not the focus of the RQs of the present study. At the interpretation posttest, the more and less explicit conditions significantly outperformed the baseline condition for all and old items, but only the less explicit condition did so for new items. Two weeks after treatment, both conditions performed better than the baseline group, but no significant differences arose. With regard to production (strict or lenient), both experimental conditions significantly outperformed the baseline condition right after treatment for all types of items. Two weeks after treatment, no significant differences arose except for old items, where the less explicit group outperformed the baseline condition.
7. Analyses for production tests lenient scoring yielded very similar results to those observed for strict scoring immediately and 2 weeks after treatment, shedding no significant differences in performance.
8. A previous analysis revealed an outlier. It was included in the final analyses reported in this manuscript, as it barely changed the significance of the chi-square or its effect size when compared to the previous analysis.
9. For more information on this type of analysis, see Hardin and Hilbe (2012).
10. When comparing present findings with Morgan-Short, Heil, et al.'s study (2012), it is important to keep in mind that while participants coded for an intermediate level of processing were roughly equal in both studies (79% and 67% M-S), those coded for a deep level of processing represented 23% in their study, and only 7% in the present experiment.

REFERENCES

Baddeley, A. D. (1978). The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. *Psychological Review, 85*, 139–152.

Benati, A. (2004). The effects of structured input activities and explicit information on the acquisition of the Italian future tense. *Processing instruction: Theory, research, and commentary* (pp. 207–225). Mahwah, NJ: Erlbaum.

Bird, S. (2012). Expert knowledge, distinctiveness, and levels of processing in language learning. *Applied Psycholinguistics, 33*, 665–689.

Bowles, M. A. (2010). *The think-aloud controversy in second language research*. New York: Routledge.

Calderón, A. (2014). *Level of intake, depth of processing, and type of linguistic item in L2 development* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

Collentine, J. (1995). The development of complex syntax and mood selection abilities by intermediate-level learners of Spanish. *Hispania*, *78*, 122–135.

Collentine, J. (1998). Processing instruction and the subjunctive. *Hispania, 81*, 576–587.

Craik, F. I. M. (2002). Levels of processing: Past, present and future? *Memory, 10*, 305–318.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268.

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Dominicis, M. C., & Reynolds, J. J. (2011). *Repase y Escriba: Curso avanzado de gramática y composición* (6th ed.). Hoboken, NJ: Wiley.

Doughty, C. J. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition, 13*, 431–469.

Erlam, R. (2003). The effects of deductive and inductive instruction on the acquisition of direct object pronouns in French as a second language. *Modern Language Journal, 87*, 242–260.

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research, 9*, 147–171.

Farley, A. (2004). The relative effects of processing instruction and meaning-based output instruction. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 143–168). Mahwah, NJ: Erlbaum.

Fernández, C. (2008). Reexamining the role of explicit information in processing instruction. *Studies in Second Language Acquisition, 30*, 277–305.

Fernández, M., & Aguado, G. (2007). Algunas medidas del desarrollo típico de la morfosintaxis para la evaluación del lenguaje espontáneo de niños hispanohablantes. *Revista de Logopedia Foniatría y Audiología, 27*, 140–152.

Gardiner, J. M., Brandt, K. R., Vargha-Khadem, F., Baddeley, A., & Mishkin, M. (2006). Effects of level of processing but not of task enactment on recognition memory in a case of developmental amnesia. *Cognitive Neuropsychology*, *23*, 930–948.

Gardiner, J. M., Gregg, V. H., Mashru, R., & Thaman, M. (2001). Impact of encoding depth on awareness of perceptual effects in recognition memory. *Memory & Cognition, 29*, 433–440.

Gardiner, J. M., Java, R. I., & Richardson-Klavehn, A. (1996). How level of processing really influences awareness in recognition memory. *Canadian Journal of Experimental Psychology, 50*, 114–122.

Gass, S., Svetics, I., & Lemelin, S. (2003). Differential effects of attention. *Language Learning, 53*, 497–545.

Greenslade, T. A., Bouden, L., & Sanz, C. (1999). Attending to form and content in processing L2 reading texts. *Spanish Applied Linguistics, 3*, 65–90.

Hardin, J. W., & Hilbe, J. M. (2012). *Generalized linear models and extensions* (3rd ed.). College Station, TX: Stata Press.

Herron, C., & Tomasello, M. (1992). Acquiring grammatical structures by guided induction. *French Review, 65*, 708–718.

Hsieh, H., Moreno, N., & Leow, R. P. (2015). Awareness, type of medium, and L2 development: Revisiting Hsieh (2008). In R. P. Leow, L. Cerezo, & M. Baralt (Eds.), *Technology and L2 learning: A psycholinguistic approach*. Berlin: de Gruyter

Hsieh, H.-C. (2008). The effects of type of exposure and type of post-exposure task on L2 development. *Journal of Foreign Language Instruction, 2*, 117–138.

Hulstijn, J. H. (1997). Second language acquisition research in the laboratory. *Studies in Second Language Acquisition, 19*, 131–143.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*, 539–558.

Java, R. I. (1994). States of awareness following word stem completion. *European Journal of Cognitive Psychology, 6*, 77–92.

Karayianni, I., & Gardiner, J. M. (2003). Transferring voice effects in recognition memory from remembering to knowing. *Memory & Cognition, 31*, 1052–1059.

Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research, 12*, 365–386.

Konstantinou, I., & Gardiner, J. M. (2005). Conscious control and memory awareness when recognizing famous faces. *Memory, 13*, 449–457.

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review, 59*, 567–587.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*, 1–26.

Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning, 47*, 467–505.

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. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. London: Routledge.

Leow, R. P., Hsieh, H. C., & Moreno, N. (2008). Attention to form and meaning revisited. *Language Learning, 58*, 665–695.

Leow, R. P., & Mercer, J. D. (2015). Depth of processing in L2 learning: Theory, research, and pedagogy. *Journal of Spanish Language Teaching, 2*, 1–14.

López Ornat, S., Fernández, A., Gallo, P., & Mariscal, S. (1994). *La adquisición de la lengua española*. Madrid: Siglo XXI.

Martínez, A. (2010). *Experiences of remembering and knowing in SLA, L2 development, and text comprehension: A study of levels of awareness, type of glossing, and type of linguistic item* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

McLaughlin, B. (1987). *Second-language acquisition in childhood: School-age children* (Vol. 2). Mahwah, NJ: Erlbaum.

Morgan-Short, K., Finger, I., Grey, S., & Ullman, M. (2012). Second language processing shows increased native-like neural responses after months of no exposure. *PLOS ONE, 7*, 1–18.

Morgan-Short, K., Heil, J., Botero-Moriarty, A., & Ebert, S. (2012). Allocation of attention to second language form and meaning: Issues of think alouds and depth of processing. *Studies in Second Language Acquisition, 34*, 659–685.

Nelson, T. O. (1977). Repetition and levels of processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 151–171.

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

Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing, 10*, 277–303.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition, 21*, 89–102.

Révész, A., & Brunfaut, T. (2013). Task characteristics of task input and difficulty in second language listening comprehension. *Studies in Second Language Acquisition, 35*, 31–65.

Richardson-Klavehn, A., & Gardiner, J. M. (1995). Retrieval volition and memorial awareness in stem completion: An empirical analysis. *Psychological Research, 57*, 166–178.

Robinson, P. (1995). Attention, memory, and the noticing hypothesis. *Language Learning, 45*, 283–331.

Robinson, P. (1996). *Consciousness, rules, and instructed second language acquisition*. New York: Peter Lang.

Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition, 19*, 223–247.

Rosa, E. M., & Leow, R. P. (2004a). Awareness, different learning conditions, and second language development. *Applied Psycholinguistics, 25*, 269–292.

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

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press

Rosa, E. M., & O'Neill, M. (1999). Explicitness, intake, and the issue of awareness. *Studies in Second Language Acquisition, 21*, 511–556.

Sanz, C., Lin, H., Lado, B., Stafford, C., & Bowden, H. (2014). One size fits all? Learning conditions and working memory capacity in ab initio language development. *Applied Linguistics*. Advance online publication.

Schmidt, R. (1993). *Attention and awareness in foreign language learning*. Honolulu, HI: University of Hawaii Press.

Shaffer, C. (1989). A comparison of inductive and deductive approaches to teaching foreign languages. *Modern Language Journal, 73*, 395–403.

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

Stafford, C., Bowden, H. W., & Sanz, C. (2012). Optimizing language instruction: Matters of explicitness, practice and cue learning. *Language Learning, 62*, 741–768.

Tagarelli, K. M., Borges Mota, M., & Rebuschat, P. (2014). Working memory, learning conditions, and the acquisition of L2 syntax. In W. Zhisheng, M. Borges Mota, & A. McNeill (Eds.), *Working memory in second language acquisition and processing: Theory, research and commentary*. Bristol: Multilingual Matters.

Terrell, T. D., Baycroft, B., & Perrone, C. (1987). The subjunctive in Spanish interlanguage: Accuracy and comprehensibility. In B. Van Patten, T. D. Dvorak, & J. F. Lee (Eds.), *Foreign language learning: A research perspective* (pp. 19–31). Cambridge: Cambridge University Press.

Timberlake, A. (2007). Aspect, tense, mood. In T. Shopen (Ed.), *Language typology and syntactic description: Vol. 3. Grammatical categories and the lexicon* (2nd ed., pp. 280–333). Cambridge: Cambridge University Press.

Toth, J. P. (1996). Conceptual automaticity in recognition memory: Levels of processing effects on familiarity. *Canadian Journal of Experimental Psychology, 50*, 123–138.

VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition, 12*, 287–301.

VanPatten, B. (1994). Evaluating the role of consciousness in second language acquisition: Terms, linguistic features, and methodology. *Association Internationale de Linguistique Appliquée, Annual Review, 11*, 27–36.

VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Mahwah, NJ: Erlbaum.

VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15*, 225–243.

VanPatten, B., & Jegerski, J. (2011). L2 processing and parsing: The issues. In B. VanPatten & J. Jegerski (Eds.), *Second language processing and parsing* (pp. 3–23). Amsterdam: John Benjamins.

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

https://doi.org/10.1017/S0142716416000308 Published online by Cambridge University Press