

ALLOCATION OF ATTENTION TO SECOND LANGUAGE FORM AND MEANING

Issues of Think-Alouds and Depth of Processing

Kara Morgan-Short, Jeanne Heil, Andrea Botero-Moriarty,
and Shane Ebert

University of Illinois at Chicago

The aims of the present study were twofold. The study addressed the issues of simultaneous attention to form and meaning in second language (L2) written input and reactivity of think-alouds. Specifically, the study examined the comprehension of L2 learners of Spanish who either attended to lexical or grammatical forms while reading for meaning or read for meaning alone. Learners completed these tasks while either thinking aloud or not. Results indicated only a minimal effect for thinking aloud that did not appear to compromise the internal validity of the study. Additionally, results showed that attending to grammatical or lexical form while reading for meaning did not affect comprehension. Indeed, learners who processed these forms more deeply evidenced greater comprehension.

We would like to gratefully acknowledge Bryan Koronkiewicz for help with data collection and processing; Charity Anne Caldwell at the University of Illinois at Chicago Language and Culture Learning Center for assistance with the technical setup for the study; Ron Leow for providing the full materials of the Leow, Hsieh, and Moreno (2008) study; members of the fall 2009 Theories of Second Language Acquisition course at the University of Illinois at Chicago for comments on the design and preliminary results of the study; and the anonymous *SSLA* reviewers for their keen insights. The results reported herein were presented at the 2010 Second Language Research Forum. Any remaining errors are fully our own.

Address correspondence to: Kara Morgan-Short, Department of Hispanic and Italian Studies, The University of Illinois at Chicago, 1706 UH, MC-315, 601 S. Morgan St., Chicago, IL 60607; e-mail: karams@uic.edu.

These findings are considered in light of methodological issues and the larger issue of simultaneous attention to form and meaning in a L2.

The role of attention has become a central issue to a wide variety of theoretical and applied perspectives in the field of SLA. In regard to theoretical perspectives, most if not all approaches to second language (L2) acquisition posit some role for attention whether it be attention at the level of noticing (Schmidt, 2001) or attention at the level of detection (Tomlin & Villa, 1994). In regard to applied perspectives, several types of L2 instruction have been developed with the goal of directing learners' attention to linguistic form, meaning, or both. Research suggests that directing learners' attention to both form and meaning may be more effective for L2 development than instruction that directs learners' attention to form or meaning alone (Norris & Ortega, 2000). Given the theoretical interests and the applied implications related to the role of attention in L2 acquisition, research that elucidates this issue clearly merits further examination.

One aspect of attention that has been examined by SLA research is how learners, particularly those at lower levels of proficiency, allocate attention to L2 aural and written input. This line of research has primarily been motivated by VanPatten's primacy of meaning principle, which states that "learners process input for meaning before they process it for form" (2004, p. 14). To test this principle, VanPatten (1990) investigated whether learners could successfully attend to both form and meaning simultaneously when presented with aural L2 input. The results of this seminal study indicated that paying attention to grammatical form while listening to a passage for meaning impaired L2 learners' ability to comprehend the passage.

VanPatten (1990) has been extended through a series of conceptual replication studies that have examined simultaneous attention to form and meaning in the written mode as well as in a different L2 (Greenslade, Bouden, & Sanz, 1999; Leow, Hsieh, & Moreno, 2008; Wong, 2001). Most recently, Leow et al. examined this issue and found that paying attention to grammatical form did not negatively impact comprehension of written L2 input. Additionally, Leow et al. extended this line of research in a new direction by exploring whether processing forms at deeper levels might have an effect on comprehension. However, they were not able to offer conclusions related to this issue due to the low number of participants who evidenced deeper levels of processing. Thus the way depth of processing may affect comprehension of a written text when L2 learners

pay attention to both form and meaning simultaneously is an open question.

As part of their study design, Leow et al. (2008) collected concurrent, nonmetalinguistic verbal protocols, which are referred to here as think-alouds (TAs). Think-alouds ask learners to verbalize their thoughts aloud as they perform a task without providing any metalinguistic explanation or justification (Ericsson & Simon, 1993). The TAs in Leow et al. were collected (a) to assess whether learners had indeed paid attention to—that is, processed—both form and meaning as directed and (b) to provide qualitative evidence about learners' processing strategies. The use of TAs in L2 research—and in research more generally—however, is not uncontroversial (see Bowles, 2010, for a detailed discussion of this issue). A clear benefit of TAs is that they can shed light on cognitive processes involved in a task, thereby increasing the internal validity of research findings (Ericsson & Simon, 1993). A possible problem of TAs, however, is their potential reactivity; that is, for some tasks TAs may actually alter the cognitive processes involved in the task. In these cases, the internal validity of the research is clearly compromised. Although TAs have been shown to be nonreactive for certain written L2 comprehension tasks (Bowles, 2008; Bowles & Leow, 2005; Leow & Morgan-Short, 2004), their potential reactivity has not been examined with tasks that direct L2 learners to pay attention to a particular lexical or grammatical form in the input while engaging in a written comprehension task. Thus, although the results from Leow et al. may potentially provide further insight into the allocation of attentional resources to L2 input, whether reactivity played a role in this study specifically and whether it plays a role more generally in tasks in which L2 learners pay attention to form and meaning simultaneously remains to be investigated.

Motivated by the line of research addressing the allocation of attention to form and meaning in L2 input (e.g., Greenslade et al., 1999; Leow et al., 2008; VanPatten, 1990; Wong, 2001), the aim of the current study, a conceptual replication based on Leow et al., is twofold. First, the study aims to replicate the findings of Leow et al. to provide further evidence regarding whether simultaneous attention to form and meaning interferes with L2 comprehension as well as whether there may be a relationship between written L2 comprehension and depth of processing. Second, the study probes the potential reactivity of the use of TAs in a L2 task that asks learners to direct their attention to form while reading a passage for meaning. The following Review of the Literature section further motivates the need to address these two issues by reexamining the extant literature relevant to (a) simultaneous attention to form and meaning in a L2 and (b) the issue of reactivity of TAs.

REVIEW OF THE LITERATURE

Attention to Form and Meaning

To date, four published studies have examined the allocation of attention to both form and meaning in a L2: VanPatten (1990), Greenslade et al. (1999), Wong (2001), and Leow et al. (2008). In his seminal study, VanPatten tested whether L2 learners have difficulty attending to form and meaning simultaneously in aural L2 input. VanPatten adopted the position from cognitive psychology that humans have a limited capacity for processing information (e.g., McLaughlin, 1987; McLaughlin, Rossman, & McLeod, 1983) and argued that L2 learners, particularly those at lower levels of proficiency, would have difficulty processing both form and meaning in aural L2 input. He predicted that asking L2 learners to pay attention to grammatical forms (which did not carry much meaning) relative to lexical forms (which were meaning bearing) would negatively affect their ability to get meaning from the input. This prediction was tested with first-semester, fourth-semester, and third-year college Spanish students who listened to a 275-word passage in their L2 (Spanish) under one of three experimental conditions or a control condition. In the experimental conditions, participants were asked to listen to the passage for content and to note with a check mark on a blank sheet of paper each occurrence of a target L2 form, which was either (a) the lexical item *inflación*, “inflation”; (b) the definite article *la*, “the”; or (c) the Spanish verbal morpheme *-n*, which indicated that the subject of a verb was third-person plural. To ensure that participants did indeed listen for these forms, participants in the experimental groups were included in the analysis only if they had provided check marks for 60% or more of the target forms. Participants in the control condition were asked to listen to the passage for content only. VanPatten hypothesized that paying attention to *la* and *-n*, grammatical forms that do not carry a lot of meaning, would interfere with participants’ ability to comprehend the passage, whereas listening to the meaningful lexical item *inflación* would not. Following the listening task, participants were asked to write down in English everything they could remember from the passage. These written recalls were later scored using an idea unit analysis as a measure of comprehension.

Results from VanPatten (1990) suggested that attending to the lexical item did not negatively affect comprehension as compared to the control, whereas attending to the grammatical items did—although the results slightly differed depending on the semester level of the participant. More specifically for all levels, participants who attended to the lexical item comprehended the passage as well as participants in the control condition. In contrast, participants at all levels who attended to the

definite article *la* did not comprehend the passage as well as control participants. In a similar manner, participants who attended to the verbal morpheme *-n* did not comprehend the passage as well as the control group, although this result held only for participants at the first-semester and third-year levels.¹ Overall, VanPatten interpreted these findings as evidence for the claim that attention to relatively meaningless (noncommunicative) grammatical items interferes with comprehension, whereas attention to meaningful lexical items does not. In other words, the results of this study suggest that L2 learners in the first years of learning do indeed have difficulty paying attention simultaneously to form and meaning when the form itself conveys relatively little meaning.

A series of studies has replicated and extended the results of VanPatten (1990). First, Greenslade et al. (1999) replicated the study in the written mode using the same Spanish passage, target forms, and tasks with L2 learners in their third semester of Spanish. Because this study was administered in the written mode and not the aural mode, participants in the experimental condition were asked to circle the target forms in the written passage as they read the passage for comprehension. Like VanPatten, Greenslade et al. found that participants who paid attention to the lexical item *inflación* comprehended the passage as well as participants in the control group, who read for content only. Participants who paid attention to the grammatical forms *la* and *-n*, however, showed impaired comprehension compared to the control group. Thus the results of Greenslade et al.'s study of simultaneous attention to form and meaning in the written mode were largely in line with the results from VanPatten's study in the aural mode. However, the authors did note that differences in the overall comprehension scores between the two studies suggest that the mode of presentation (i.e., aural vs. written) does influence how learners process forms.

In a study with L1 French learners of English, Wong (2001) examined simultaneous attention to form and meaning in both the written and the aural modes in low-intermediate L2 learners. In this study, a control group was compared to groups of participants who attended to either the lexical item, *inflation*, or the definite article in English, *the*. Apart from the target language and the omission of a second grammatical item, Wong closely followed the methods used in VanPatten (1990). In regard to the aurally presented passage, the results from Wong's study were largely consistent with VanPatten in that the group that attended to the lexical form comprehended the passage at a similar level as the control group, whereas the group that attended to the grammatical form did not comprehend the passage as well as the control group. For the written modality, however, the pattern of results was quite different: All groups comprehended the passage at a similar level. On the basis of these results, Wong suggested that learners' attentional resources

may be constrained differently in the aural and written modes. More specifically, a learner's limited attentional capacity may not be constrained to the same extent when processing written input as when processing aural input. To explore such constraints, or lack thereof, Wong suggested that future research consider the use of TAs to provide further insight into the relationship between attention to meaning and form when processing input.

Leow et al. (2008) continued this strand of inquiry in the written mode, making adjustments in the methods to address some methodological issues with the previous studies. The changes included (a) the collection of TAs as a means of determining whether participants attempted to process to both form and meaning as instructed; (b) the use of a 10-item written multiple-choice test as the measure of comprehension instead of the free recall method; (c) the use of the Spanish word *sol*, "sun," instead of *inflación* as the lexical item to control for saliency differences between the lexical and grammatical items; (d) the inclusion of a new grammatical form, *lo*, a direct object clitic pronoun meaning "him" or "it," which was claimed to have a higher communicative value than the definite article and the verbal morpheme; and (e) a more even distribution of the target forms throughout a new reading passage that had been modified from an authentic Spanish article. Participants in this study, L2 learners of Spanish enrolled in a second-semester university Spanish course, read a 358-word Spanish passage, paying attention to meaning alone (control group) or to both form and meaning (experimental groups). They subsequently completed the written comprehension test. Participants in the experimental groups were included in the analysis if they had circled at least 60% of the occurrences of their particular target form. Similar to the results of the written passage in Wong (2001), Leow et al. found no differences in the level of comprehension between the experimental and control conditions. Comparing these results with those from studies within this strand of research that had examined aural input, Leow et al. suggested that different cognitive constraints for processing aural and written input could provide one possible explanation for the differing sets of results.

The coding of the TA data in Leow et al. (2008), however, revealed another possible explanation for the results. The TA data revealed an apparent issue related to the depth or level of processing (Craik & Lockhart, 1972) in that different learners processed the target forms at different levels (e.g., merely circling the form vs. interpreting the form). To explore potential effects related to how deeply the target form was processed, TA data for each participant was categorized into one of three processing levels: (a) Level 1 participants had merely circled the target forms, (b) Level 2 participants had provided some additional report of attending to the target form, and (c) Level 3 participants had

interpreted or translated the target form. On the basis of VanPatten's primacy of meaning principle, the authors hypothesized that increased depth of processing would be associated with reduced comprehension. The results, however, did not show a relationship between the level of processing and comprehension, which was due at least in part to the fact that the number of participants at each level for each form was small, which precluded any statistical analysis or strong conclusions about the relationship or lack thereof. Thus the authors of the study suggested that future research address the issue of the depth of processing of form.

In sum, results from studies that have examined the allocation of attention to form and meaning suggest that, for aurally presented L2 input, learners may have difficulty attending simultaneously to form and meaning (VanPatten, 1990; Wong, 2001). Although some evidence exists that this is also true for written L2 input (Greenslade et al., 1999), other evidence suggests that this attentional constraint is not relevant in the written mode (Leow et al., 2008; Wong, 2001). Even though the mode of processing is likely to play some role in these differing sets of results, as suggested by Leow and colleagues, whether and how deeply learners processed both form and meaning are other factors that may account for the results and should be explored further through empirical investigation.

The Issue of Reactivity of Think-Alouds

One of the contributions of Leow et al. (2008) to this line of research was the collection of TAs, which were used to (a) determine whether participants included in the analysis were representative of their condition (i.e., attending to form and meaning simultaneously as directed) and (b) provide insights into participants' cognitive processes. The use of TAs in L2 research, however, has itself been the subject of considerable debate (see Bowles, 2010, for a detailed consideration of this issue in cognitive psychology and L2 acquisition research). In regard to concurrent TAs (i.e., protocols collected while participants complete a task), the issue of contention is the potential reactivity of thinking aloud, or, in other words, whether the act of thinking aloud alters the cognitive processes involved in performing the task. Thus, although the results from Leow et al. may potentially provide further insight into the allocation of attentional resources, whether reactivity played a role in this study remains to be empirically investigated.

Although Leow et al. (2008) did not include a non-think-aloud (NTA) group to directly examine potential reactivity effects, the researchers

did address the appropriateness of using TAs in their study. They noted that three L2 studies (i.e., Bowles, 2008; Bowles & Leow, 2005; Leow & Morgan-Short, 2004) have examined the issue of reactivity of TAs in written tasks and “have reported that metalinguistic and/or non-metalinguistic verbalization during text processing did not significantly affect learners’ subsequent performance when compared to a control group, except for latency” (p. 676). However, more recent work is relevant both to the general issue of reactivity of TAs and to the specific use of TAs in Leow et al. Bowles (2010) provided a thorough qualitative and quantitative review of the issue of reactivity of TAs in a L2. In her review, she identified four studies that examined reactivity when participants’ task was to read a passage, which is one of the aspects of the task used in studies of simultaneous attention to form and meaning. In regard to accuracy, two of these studies reported nonreactivity as assessed on posttask performance (Bowles & Leow, 2005, and Leow & Morgan-Short, 2004, as cited by Bowles, 2010), whereas the other two studies did find reactivity (Polio & Wang, 2005, as well as Rossomondo, 2007, as cited by Bowles, 2010). The results of Bowles’s quantitative, exploratory meta-analysis indicated that nonmetalinguistic TAs had a small effect ($d = 0.21$) on posttask assessments of comprehension, although this effect should not be considered reliable, given that zero fell within the range of the 95% confidence intervals. More generally, on the basis of the overall results of the exploratory meta-analysis across different L2 tasks and posttask assessments, Bowles concluded that “thinking aloud while completing a verbal task has a small effect on post-task performance” (2010, p. 110)—even though effect sizes varied with different tasks. It is also important to note that in her review of the issue of reactivity in the larger cognitive psychology literature, Bowles found that tasks that were more cognitively demanding or complex, as a dual task might be, tended to have reactive effects. In conclusion, Bowles stressed the importance of including a NTA group to determine whether thinking aloud during a particular task affects performance. Given Bowles’s findings and conclusions about TAs in L2 research, and because attending to both form and meaning—as operationalized by reading for comprehension and circling a target form—is arguably a dual task, it seems prudent to examine data from a NTA group to probe whether the TA protocol in Leow et al.’s study design might have led to reactive effects.

Motivation and Research Questions

Considering the dual nature of the questions left open by Leow et al. (2008) and by the literature on simultaneous attention to form and

meaning when processing L2 input more generally (Greenslade et al., 1999; VanPatten, 1990; Wong, 2001), the aims of the present study are twofold: (a) to replicate and extend the results of Leow et al. and (b) to examine whether having participants think aloud while attending to form and meaning might have reactive effects. To address these aims without introducing additional potentially confounding variables, the study was designed as a conceptual replication of Leow et al.

Researchers have pointed out that replication is an essential element of the scientific method and is crucial to distinguish spurious from real results (Polio & Gass, 1997) and to ensure that L2 acquisition research is valid and reliable (Valdman, 1993). Thus the current study is designed to serve two purposes toward this end. First, it addresses the internal validity of Leow et al. (2008) by examining whether the TAs administered in that study might have affected the processes that they were expected to reveal. Second, it addresses the external validity of Leow et al. by aiming to replicate the results regarding the lack of an effect of simultaneous attention to form and meaning on written comprehension in a different group of participants, and by examining the issue of depth of processing—for which data from Leow et al. were inconclusive. More specifically, the current study addresses the following research questions and hypotheses:

1. Does thinking aloud affect participants' comprehension of that passage when they are asked to pay attention to form and meaning simultaneously?

Based on the conclusions from Bowles (2010) that thinking aloud during L2 tasks leads to a small positive effect and that, in cognitive psychology, reactivity is more common in complex tasks, this study adopts the hypothesis that thinking aloud may have a reactive effect when participants are reading a passage for comprehension and simultaneously paying attention to a particular form in the passage.

2. Does simultaneous attention to form and meaning in L2 written input affect comprehension?

The null hypothesis (i.e., that there will not be differences in comprehension among different attentional conditions) is posited for the second research question. Previous research addressing this question in the written mode has, in large part, not found effects of attention to form and meaning as compared to attention to meaning alone. Although Greenslade et al. (1999) found negative effects for attention to grammatical forms, Leow et al. (2008) and Wong (2001) did not find any effects for grammatical or lexical forms compared to the control. Given that the design of the current study closely follows that of Leow et al., it is expected that their findings will be replicated.

3. Is there a relationship between the overall comprehension of a passage and the depth of processing of a particular form?

The null hypothesis (i.e., that there is no relationship between comprehension and depth of processing a form) is adopted for this question. Although

a logical prediction on the basis of VanPatten's primacy of meaning principle may be that deeper levels of processing interfere with written comprehension, empirical support has not yet been provided for this position. Evidence provided by Leow et al. (2008) did not support or refute this prediction because the depth of processing reported in that study appeared to be relatively low and the data set was not large enough to statistically probe any relationship, or lack thereof, between the level of processing and comprehension.

To address these research questions and hypotheses, the current study closely follows the research design of Leow et al. (2008) with the following modifications: (a) a NTA group was added to assess whether thinking aloud while attending to form and meaning is reactive and (b) the number of participants was increased to facilitate examination of any relationship between depth of processing and comprehension. Apart from these modifications, all other aspects of the design remain as faithful as possible to the design of Leow et al. to facilitate comparability of the results between the studies.

METHODS

Participants

Participants for this study originally consisted of 410 college students enrolled in a third-semester Spanish course at a Midwestern university (approximately half male and half female; M age = 20.7 years). This course was part of a four-semester basic Spanish language program that had a communicative focus. The first three semesters of the program were based on the text *Sol y Viento* (VanPatten, Leeser, Keating, & Román-Mendoza, 2005). Data were collected using intact classes over the fall and spring semesters, with 139 students (from 12 classes) and 271 students (from 15 classes) participating each semester, respectively. The study occurred during the seventh week of each semester. Because the syllabus was standardized across classes and across semesters, all students had received very similar instruction in the basic Spanish language program prior to participating in the study. Half of the participants were assigned to a TA group and half were assigned to a NTA group (see Procedures).

Although participants in the Leow et al. (2008) study were second-semester students, this study was run with third-semester students because of the heavy use of the preterit and imperfect verb forms (i.e., two kinds of past tense in Spanish) in the comprehension passage. In the basic Spanish language program at the Midwestern university where this study took place, students were not exposed to both of these forms until the end of the second semester. Thus the most appropriate level

of administration of the study was with third-semester, rather than second-semester, students. Although an independent measure of proficiency is not reported by the current study or by Leow et al. (2008), it appeared likely that participants in this study were not more advanced than participants from Leow et al. A post hoc examination of the syllabi and contents of their textbooks confirmed this assumption: Students had completed similar proportions of their respective Spanish textbooks and had been instructed on a similar range of structures.²

To be included in analyses, participants in the experimental groups were required to demonstrate that they had paid attention to a target form while reading a Spanish passage for comprehension by having circled 60% or more of the target forms in the passage. On the basis of this criterion, 45 participants were excluded from analysis. An additional 4 participants were excluded from analysis because they did not complete the tasks of the study as instructed (e.g., they used the passage while answering comprehension questions). Of the 49 participants who were excluded, 31 had been assigned to the TA group (attrition rate = 15.12%), and 18 had been assigned to the NTA group (attrition rate = 8.78%). A chi-square analysis of the number of excluded participants per TA group revealed a significant relationship between TA group and attrition, $\chi^2(1, n = 410) = 3.917, p = .048$, such that participants from the TA group were more likely to be excluded than participants from the NTA group. Overall, the attrition rate was 12% (15.12% for the TA group, and 8.78% for the NTA group), and the number of participants included in the analyses was 361.

Materials

The materials for this study were adopted from Leow et al. (2008). There were four experimental conditions and one control condition. The experimental conditions were based on participants being instructed to circle a particular target form while reading a Spanish passage for comprehension. In the control group, participants only read the passage for comprehension. The target forms used in the experimental conditions included one lexical item and three grammatical items. The lexical item was the word *sol*, “sun.” This lexical item had been chosen by Leow and colleagues in part because it was monosyllabic, as were all the grammatical items, and thus was assumed not to be more salient than the grammatical items. The grammatical items included the definite article *la*, “the”; *-n*, the verbal morpheme marking the subject of the verb as third person plural; and *lo*, a clitic pronoun meaning “him” or “it.” The first two grammatical items had been included in

previous studies (Greenslade et al., 1999; VanPatten, 1990), and the third grammatical item was introduced by Leow et al. because it “carries both meaning and grammatical function in the input” (p. 674).

As with the target forms, the reading passage and written multiple-choice comprehension tests were adopted directly from Leow et al. (2008)—the reader is referred to Appendices A and B in Leow et al. for materials. Leow et al. reported that the reading passage was a modified version of an authentic article about Aztecs written in Spanish by Rangel Montemayor, which had been shortened to 354 words.³ Additionally, the four target forms (*sol*, *lo*, *la*, and *-n*) had been balanced in such a way that each form occurred 10 times and was equally distributed among the paragraphs of the passage. The written comprehension test consisted of 10 multiple-choice items that had been designed to closely follow “the corresponding number of targeted forms in the text” (Leow et al., 2008, p. 667). The test was entirely in English, the participants’ L1, which ensured that comprehension of the passage and not of the questions themselves was being tested. Finally, there was a sample practice TA activity, also adopted from Leow et al. The activity was a problem-solving activity that required participants to think aloud while they calculated the amount of money needed to purchase a short list of grocery items.

Procedure

At the start of the study, intact third-semester Spanish classes (in total, 27 classes over two semesters) were pseudorandomly assigned to be part of the TA group or the NTA group, resulting in 205 participants being assigned to each group. The pseudorandomization controlled for the hour of the class (e.g., 9:00 am), such that various class hours were assigned to the TA and NTA groups as equally as possible. For the current study, there were multiple reasons that motivated the inclusion of a TA group. First, to address the external validity of Leow et al. (2008) results through replication, the current study included a TA group so that the current study’s results would be as comparable as possible to those of the original study. Second, given that one of the primary aims of the current study was to test the potential reactivity of TAs administered when learners were directed to pay attention to both form and meaning, it was necessary to collect data from TA and NTA groups to compare performance under the two conditions. Third, as pointed out by Leow et al., the use of TAs provides an opportunity for researchers to confirm whether participants performed the experiment as directed and whether they are representative of their assigned experimental condition. Finally, assuming that

thinking aloud does not interfere with the cognitive processes used by the learner, analysis of the TA data has the potential to reveal which cognitive processes are being engaged as learners complete the task (Ericsson & Simon, 1993). It is important to note that these last two reasons also served as the motivation for administering TAs in the Leow et al. study.

On the day of the study, the participants in the NTA classes reported to their classrooms, whereas participants in the TA classes reported to a multimedia classroom equipped with laptops and headphones with built-in microphones for each participant. Apart from a few additional steps for the TA group related to thinking aloud, the TA and NTA groups both followed the same procedure. All participants were asked to provide informed consent. Next, participants in the TA group were provided with instructions regarding how to set up their headphones and microphones and how to use the recording software necessary for the study (Voice Board, accessed through the Blackboard course management system). They then completed the practice TA activity. Researchers monitored the classroom during this time to ensure that all participants were indeed thinking aloud.

Once participants were ready to begin the experimental tasks, the researchers distributed packets that contained instructions, the reading passage, and the written comprehension test. The packets had been arranged so that each consecutive participant received a packet from a different condition and so that the distribution of packets in each classroom began with a different condition. In this way, a random yet even distribution of all conditions was maximized. Once participants received their packet face down, they were provided with a short aural overview of what they would be doing and were then asked to turn over their packet. The researchers asked the participants to read the instructions on their packets, which directed participants to read the passage for comprehension and then to answer some questions related to the passage. Participants who received experimental packets were also instructed to circle all occurrences of a particular target form as they read the passage and were provided with an example sentence containing an occurrence of the target form with a circle around it to illustrate what they should do. Finally, in addition to the instructions related to the passage and comprehension tests, participants in the TA group were instructed to think aloud as they completed the packet. See the Appendix for sample instructions.

When all participants had completed reading the instructions, the researcher instructed the participants to begin, and to begin recording if they were in the TA group. Participants then read the passage according to their group (TA or NTA) and their condition (control, *sol*, *lo*, *la*, or *-n*) and completed the written comprehension test at their own pace.

Scoring

Each multiple-choice question of the comprehension test was worth 1 point. Participants received 1 point if they chose the correct answer and 0 points otherwise. Due to an error in printing, 21 of the comprehension tests did not include item 7. To ensure that the comprehension score was based on the same set of questions across participants, item 7 was excluded from analysis for all participants. Thus the final score was calculated as a percentage based on a total of 9 points.

Coding

The recordings for the TA group were coded by three researchers for two purposes: (a) to establish that participants had processed the passages for form and meaning as instructed and (b) to determine the depth of their level of processing of the target form.⁴ To establish that participants had processed the passage simultaneously for form and meaning as instructed, the researchers first assessed the TAs. If the TAs revealed that a participant either had only searched for target forms or had backtracked to reread the passage for comprehension, he or she was identified as not having paid simultaneous attention to form and meaning. These participants were coded as *to be removed*. A total of 53 participants showed evidence of backtracking in their TAs and were removed from the data set. This coding distinction was not available for the NTA group because the lack of TAs precluded an examination of whether NTA participants had engaged in the experimental tasks as instructed. After participants who had not processed both form and meaning were removed, the final data set consisted of 308 participants in the TA and NTA groups ($n = 121$ and 187 , respectively).

Next, researchers assessed the verbal protocol of the participants in the TA group to determine the depth of their level of processing of the target form. Researchers noted which target form a participant had been instructed to circle and then focused on the parts of the TAs relevant to occurrences of that target form. Following Leow et al. (2008), participants were assigned a processing level of 1, 2, or 3, on the basis of what was revealed in their TA. A level of 1 was assigned if the participant had circled no less than six of the target forms but did not show any additional evidence of processing that form in their TA. A level of 2 was assigned if participants showed evidence of minimally processing at least one target form beyond just circling it. Evidence of minimally processing a target form included (a) pronouncing the target form

(e.g., stating *todo el mundo lo visita diariamente*, “the whole world visits it daily,” where *lo* is the target form; stating *no lo hacían*, “they did not do it,” where *-n* is the target form), (b) noticeably pausing immediately before or after the target form (e.g., stating *necesitaba que . . . lo . . . alimentaran*, “it was necessary that they fed it,” where *lo* is the target form and a pause occurred before and after the form; stating *a quien honraban . . . los aztecas*, “who the Aztecs honored,” where *-n* is the target form and a pause occurred after the form), or (c) making a non-metalinguistic comment about the form (e.g., “oh, there’s another *lo*,” where *lo* is the target form; “gotta circle that *sol*,” where *sol* is the target form). Finally, a level of 3 was assigned if the participant showed evidence of processing the target beyond a minimal level. Participants were determined to be at this level if they (a) made a metalinguistic comment about the target form (e.g., “oh, yeah, this means plural” after encountering an *-n*) or (b) provided a translation of the target form (e.g., “they didn’t do it,” which is the correct translation for the phrase *no lo hacían*, where *-n* is the target form; “the Metropolitan Cathedral” for *la Catedral Metropolitana*, where *la* is the target form; “they admire him a lot,” which is the correct translation for the phrase *lo admiran mucho*, where *lo* is the target form).⁵

At the onset of coding each semester, a norming session was conducted to ensure that each of the researchers was consistent in their use of the coding protocol. First, the researchers developed or revisited the coding protocol, as described previously. They then coded the same five TAs from different conditions. For each semester, a comparison of the coding results revealed that the researchers’ codes for each of the five sample TAs were 100% consistent. Thus, the researchers divided the remaining TAs and coded them individually.

Analyses

To probe the first two research questions (i.e., the potential reactive effects of thinking aloud on participants’ comprehension of the Spanish passage and the potential effects of simultaneous attention to form and meaning on comprehension), the study’s two independent variables, group (i.e., TA, NTA) and condition (control, *sol*, *lo*, *la*, and *-n*), were entered into a two-way ANOVA, which could reveal any main effect of the independent variables or any interaction between the two variables. To address the third research question (i.e., the relationship between comprehension and depth of processing), a Pearson’s two-tailed correlation was calculated between participants’ comprehension scores and their level of processing. The alpha level for all analyses was set at .05.

RESULTS

In regard to the first and second research questions, participants' scores on the comprehension test were examined. Overall the scores in the NTA group ranged from 32% to 35% accuracy on the comprehension test, whereas the scores in the TA group ranged from 24% to 35% (see Table 1).⁶ As observed in Figure 1, accuracy in the TA group was somewhat lower in general and more variable between conditions than accuracy in the NTA group. Before running parametric statistics to probe any significant effects or interactions in the results, the data were checked to ensure that assumptions of ANOVAs were met: Visual inspection of the distribution of the data and an acceptable skewness value ($< \text{absolute } 1$) indicated that comprehension scores in both the TA and NTA groups were normally distributed. Additionally, Levene's test of equality of error variances indicated that the variance between groups was not statistically different, $F(9, 298) = 1.62, p = .11$. As all of the assumptions were met, the planned two-way ANOVA was run and revealed a main effect of TA, $F(1, 298) = 3.85, p = .05, \eta^2 = .01$; no effect for condition, $F(4, 298) = 1.69, p = .15, \eta^2 = .02$; and no interaction between TA and condition, $F(4, 298) = 0.86, p = .49, \eta^2 = .01$. In regard to the factor group, these results seem to indicate that there was a difference in performance between the NTA and TA groups. It is necessary to note, however, that the amount of variance in the data that can be explained by group assignment (only 1%) is minimal, and that the

Table 1. Accuracy scores on comprehension test by TA group and condition

Condition	<i>M</i>	<i>SD</i>
A. NTA Group		
Control (40)	0.32	0.17
<i>sol</i> (42)	0.32	0.15
<i>lo</i> (37)	0.35	0.18
<i>la</i> (34)	0.35	0.13
<i>-n</i> (34)	0.34	0.20
Total (187)	0.34	0.17
B. TA Group		
Control (42)	0.24	0.14
<i>sol</i> (27)	0.30	0.14
<i>lo</i> (18)	0.27	0.12
<i>la</i> (18)	0.35	0.20
<i>-n</i> (16)	0.33	0.12
Total (121)	0.29	0.15

Note. Numbers in parentheses in the first column represent the total number of participants per condition.

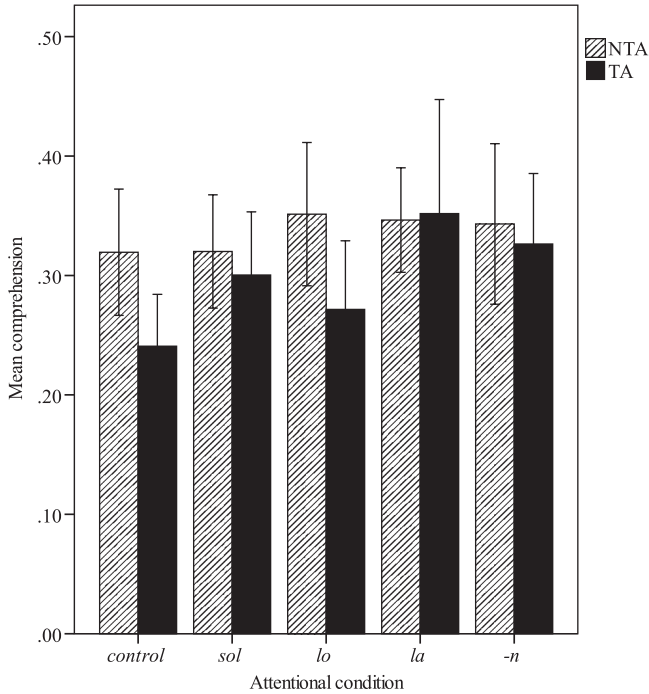


Figure 1. Comprehension scores per TA group per attentional condition.

η^2 value of .01 does not meet the .04 criterion to be considered a “recommended minimum effect size representing a ‘practically’ significant effect for social science data” (Ferguson, 2009, p. 533). In other words, although the p value for group was equal to .05, the size of the effect was minimal. In regard to condition, the results of the ANOVA indicated that paying attention simultaneously to form and meaning did not affect comprehension.

Finally, to address the third research question, participants’ scores on the comprehension test were examined in light of their depth of processing of the target form. As in Leow et al.’s (2008) data, the TA group had been coded for depth of processing on a 1–3 scale, with 1 as the lowest level and 3 as the highest level. Overall, 8 participants (10.13%) were coded at Level 1, 53 participants (67.09%) were coded at Level 2, and 18 participants (22.78%) were coded at Level 3. This general distribution seemed to hold for all attentional conditions, with the majority of participants being coded at Level 2 (see Table 2). After confirming that the data were normally distributed on the basis of visual inspection and an acceptable skewness value (< absolute 1), a Pearson’s two-tailed correlation was run between participants’ level of processing and their comprehension score. The analysis revealed a reliable

Table 2. Accuracy scores on comprehension test per condition by depth of processing

Condition	Average Comprehension Score		
	Level 1	Level 2	Level 3
Control	N/A	N/A	N/A
<i>sol</i> (27)	0.22 (2)	0.29 (19)	0.37 (6)
<i>lo</i> (18)	0.22 (3)	0.27 (13)	0.39 (2)
<i>la</i> (18)	0.39 (2)	0.27 (11)	0.51 (5)
<i>-n</i> (16)	0.22 (1)	0.33 (10)	0.33 (5)

Note. Numbers in parentheses represent the number of participants per condition and level.

positive correlation between the variables ($r = .31, p < .01$), with deeper levels of processing being related to higher comprehension scores. In terms of the size of the effect, the r value of .31 is above the .20 r value that is considered the recommended minimum effect size representing a significant effect (Ferguson, 2009), and it can be interpreted as a medium-sized effect (Cohen, 1988; Valentine & Cooper, 2003).

DISCUSSION

To summarize the findings of the current study, in regard to the first research question, probing the reactivity of TAs, the results showed that thinking aloud had a statistically significant reactive effect when participants were reading a passage for comprehension and simultaneously paying attention to a particular form in the passage. It is important to point out, however, that the size of this effect was minimal. As for the second research question, related to the effect of paying attention to form while reading a text for comprehension, the results showed that the type of attentional condition (i.e., paying attention to either *sol*, *lo*, *la*, or *-n*) did not have a detrimental effect on adult L2 reading comprehension. Moreover, the results related to the third research question (i.e., the relationship between the depth of processing a form and L2 comprehension) demonstrated that the deeper the learner's level of processing, the better his or her comprehension score of the written text. In all, the results from the current study are largely consistent with previous research related to the issues of reactivity of TAs (e.g., Bowles, 2010) and simultaneous attention to form and meaning in the written mode (Leow et al., 2008; Wong, 2001). Additionally, the results related to depth of processing provide new insights into the question of whether processing written L2 input for both form and meaning simultaneously is detrimental to comprehension, as may be predicted by

VanPatten's primacy of meaning principle. The relationship to previous research and the more general implications of each of these results are discussed in turn.

One of the primary goals of the study was to examine whether thinking aloud would result in reactive effects when L2 learners read a passage for comprehension and attended to a target form. If TAs are reactive in this case, the results reported in Leow et al. (2008) may have been confounded by the fact that all participants in that study thought aloud while completing the task. The results from the current study found that thinking aloud led to a statistically significant effect on comprehension ($p = .05$), although the size of the effect was minimal ($\eta^2 = .01$). Given the minimal effect size, one interpretation of this result is that even though it is statistically significant, it may not be a practically significant effect (Ferguson, 2009). Larson-Hall (2010) points out that "if the effect size is quite small, then it may make sense to simply discount the findings as unimportant, even if they are statistical" (p. 114). This interpretation of the results is also consistent with the fact that no interaction between TA group and condition was evidenced. In other words, having thought aloud did not change the statistical pattern of results between groups in regard to the effect of condition in the current study. The interpretation of the effect of TAs in this study as not being a practically significant effect is arguably also in line with the results of Bowles (2010). It is important to note that the 95% confidence interval of the small effect size reported in Bowles' exploratory meta-analysis overlapped zero, which indicated that the effect size value "is not significantly different from zero" (p. 138)—that is, it is not a reliable effect. In all, the administration of TAs did not appear to compromise the internal validity of results related to simultaneous attention to form and meaning in the current study. This conclusion might also be extended to Leow et al. (2008), given the replication nature of the current study. Although these results indicate that administration of TAs did not compromise the internal validity of this study and imply the same conclusion in regard to Leow and colleagues, it remains advisable to concur with the suggestion—posited in Leow and Morgan-Short (2004) and echoed by others (e.g., Bowles)—that it is important for future studies to collect data from control NTA groups to confirm that reactivity is not a significant issue for a study's design, in particular when the design includes a task for which the reactivity of TAs has not been previously tested. Such an approach will not only be useful in confirming the internal validity of future studies but will also contribute to the larger literature related to the issue of reactivity of TAs.

Given that the study's internal validity did not appear to be compromised by the administration of TAs, it is important to return to the second primary goal of the study, which was to contribute to an understanding

of the role of attention in L2 acquisition through an examination of the effects of the allocation of simultaneous attention to form and meaning on L2 comprehension. As hypothesized, the current study replicated the findings of Leow et al. (2008): The results indicated that varying attentional conditions did not differentially affect comprehension of written input. These findings are also consistent with those related to written comprehension reported by Wong (2001) but differ from the results related to written comprehension reported by Greenslade et al. (1999) and the aural comprehension results from VanPatten (1990) and Wong.

As suggested by Leow et al. (2008) and Wong (2001), a viable explanation of the overall set of results is that successful attention to form and meaning is dependent on the modality of the passage. Three out of four studies—that is, the current study, Leow et al., and Wong—that have examined the issue of simultaneous attention to form and meaning in the written mode have found that attentional condition does not affect comprehension. In contrast, two out of two studies—that is, VanPatten (1990) and Wong—administered in the aural mode have found effects of attentional condition. Modality, however, does not completely account for all of the results in this line of research. Greenslade et al. (1999) found that attending to grammatical form did affect comprehension in the written mode. What may account for the difference between the results from Greenslade et al. as compared to the results of Leow et al., Wong, and the current study? It is necessary to first consider the difference in the passages. In particular, the Spanish passage used in Leow et al. and the current study was arguably more difficult than the Spanish passage used in Greenslade et al. Although the number of words per sentence (approximately 16) was comparable between the two passages, the passage used in the current study and in Leow et al. was almost 30% longer and contained more than two times the number of subordinate clauses headed by the relative pronoun *que* “that” than the passage used in Greenslade et al. It is more difficult to compare the Greenslade et al. passage to the Wong passage, as the Wong passage was an English passage (although it was a translation of the Spanish passage used in Greenslade et al.). Thus, whether the passages themselves interact with attentional condition in simultaneous attention to form and meaning studies remains an empirical question.

A second possible explanation for the differing results between studies conducted in the written mode is the proficiency level of the L2 learners. The learners in the control condition in Wong (2001) recalled an average of 12 out of 52 idea units from the passage ($SD = 5.6$). In Greenslade et al. (1999), learners in the control condition recalled 22.5 out of 53 idea units ($SD = 5.9$). Thus the participants in Greenslade et al. appeared to have comprehended the written passage to a greater degree

than the participants in Wong, which might indicate different levels of proficiency between the two sets of participants. Effects of simultaneous attention to form and meaning might only be detected when learners are beyond the lowest levels of proficiency, where floor effects on the comprehension assessment might be found. It is difficult to compare the proficiency in Greenslade et al. and Wong to Leow et al. (2008) and the current study, even through comparison of the comprehension scores of the control groups, because Leow et al. and this study used a different means of assessing comprehension (i.e., multiple choice vs. recall). Indeed, no study in this line of research has administered an independent measure of proficiency, making it problematic to consider an account of the differences among studies on the basis of proficiency. Although the characteristics of the passage and the participants' proficiency level may play some—as of yet unknown—role in this line of research, they do not contradict the observation that modality does seem to account for the general pattern of results among studies. In large part, it appears that learners who are directed to attend simultaneously to form and meaning are able to comprehend a written passage as well as learners who are not directed to attend to any particular form. This does not appear to be the case for aurally presented passages.

What may explain the difference between processing both form and meaning in written input versus aural input? Previous studies in this line of research have suggested that the cognitive constraints may differ when processing aural versus written L2 input (Greenslade et al., 1999; Leow et al., 2008; Wong, 2001). Wong found that comprehension overall was better in the written mode compared to the aural mode and suggested that “learners’ attentional capacity is not constrained in the same way during input processing in the aural and written modes” (p. 358). Although a strict interpretation of this claim may be that processing aural input is subject to the attentional constraints of a limited-capacity processor, whereas processing written input might be relatively free of such constraints, a more parsimonious account of the differences between the aural and written modes may be that the differing parameters of the modes drive the differing results. It is important to note that in studies with an aural mode of presentation, the timing of the presentation of the input was controlled by the researcher, whereas in studies with a written mode of presentation, participants read the text and circled the target form at their own pace. A recent review of cognitive psychology research on attention (Chun, Golomb, & Turk-Browne, 2011) suggests that attentional processes and constraints essentially work the same in all modalities, including aural and written, but that attentional constraints are more apparent in the written mode when the timing of input is controlled through rapid serial visual presentation of stimuli. It would be interesting to see future research that

compares the effects of attentional conditions on aural and written input when both types of input are presented with similar timing parameters. Such a design would make it possible to tease apart effects of mode versus timing and would be informative as to whether VanPatten's primacy of meaning principle is applicable to both aural and written input.

To further understand the role of attention in L2 comprehension, the current study also explored whether there was a relationship between the overall comprehension of a passage and depth of processing of a particular form. Leow et al. (2008) had posited that, in addition to the issue of modality, the level or depth of processing (Craik & Lockhart, 1972) may explain the lack of effects in studies that did not find an effect for the attentional condition. This conclusion was based on the observation that few participants in that study had attended to the target forms at a deeper processing level. Leow et al. suggested that attending to forms may not have sufficiently depleted learners' attentional resources so as to interfere with attention to meaning and concluded that "the nonsignificant difference in comprehension between experimental conditions might be attributed to the relatively low level of processing reported in all experimental groups in regard to the targeted form" (p. 686). In general, Leow et al. concluded that their results did not appear consistent with the prediction that deeper processing of form would negatively impact comprehension. However, they also noted that the low number of participants in each condition at each level precluded statistical analyses regarding this question.

In this study, there were sufficient data to perform a correlational analysis, which showed that deeper levels of processing were associated with increased comprehension of written input—a relationship that would not seem consistent with VanPatten's primacy of meaning principle. The relationship, however, may be consistent with predictions that could be made by the levels of processing framework in cognitive psychology (Craik, 2002; Craik & Lockhart, 1972), which makes the claim that memory for items—as might be reflected by the scores on the comprehension test in the current study—depends on the depth of processing during encoding. The framework posits that shallow processing (i.e., processing of sensory and surface aspects) is less likely to lead to memory recall than deep processing (i.e., processing of meaning and implication). The results from this study appear to reflect this prediction: Shallower levels of processing (e.g., merely circling the target form) were associated with lower comprehension scores, whereas deeper levels of processing (e.g., interpreting the target forms) were associated with increased comprehension scores. Considering the results related to both the (non)effects of attentional conditions on comprehension and the relationship between depth of processing and

comprehension, the findings of the current study suggest that simultaneous attention to form and meaning does not interfere with comprehension when the forms are processed in a meaningful manner—which in fact leads to increased comprehension.

Limitations

As with all studies, there are certain limitations that should be noted. First, although the correlational results have been discussed in terms of the levels of processing framework, there are several issues inherent to that framework that must be worked out by SLA researchers if it is to be useful in understanding the role of attention in L2 acquisition. For example, a limitation of the framework itself is the “absence of an objective index of depth of processing” (Craik, 2002, p. 308). Indeed, questions regarding the qualitative or quantitative differences between the levels will need to be addressed so that researchers know what needs to be measured before an index of processing can be established. Also, as mentioned by an anonymous *SSLA* reviewer, would the predictions of the framework change on the basis of learner proficiency? A thorough consideration of these issues might be a fruitful direction for future research.

The current study was designed to faithfully replicate Leow et al. (2008) while also examining the potential reactivity of TAs and contributing to the issue of depth of processing. In order for the results to be as comparable as possible to those of Leow et al., the researchers of the current study made the explicit decision not to manipulate or control factors that had been pointed out as limitations by the authors of Leow et al. Thus, although this study inherited the methodological advantages of that study—such as an even distribution of target forms throughout the passage—it also inherited the methodological limitations. For example, a certain amount of participants (as in Leow et al.) were eliminated from analysis due to backtracking. Although participants were directed to circle forms as they read, more explicit directions to not backtrack might have at least partially remedied this issue. As suggested in Leow et al., a computer-administered study could largely prevent this problem in future research. The issue of simultaneous attention to form and meaning could also be addressed using a different experimental paradigm (e.g., eye-tracking), which, in addition to allowing researchers to prevent and identify backtracking, might also provide additional measures related to an individual’s level of processing.

These suggestions necessarily lead toward a more controlled and perhaps narrower experimental design. However, as suggested by an

anonymous *SSLA* reviewer, further insight into the issue of attention to form and meaning may also be gained by adopting an ethnographic examination of this issue: Transcripts of verbal protocols from L2 learners reading for meaning (with no direction to focus on form) could be analyzed for any evidence that focus on form while engaged in more natural reading affects comprehension. Converging data from experimental and ethnographic approaches might prove to be particularly useful in moving the field toward a deeper understanding of the implications of simultaneous focus on form and meaning.

Another potential limitation of the current study is that the participants were third-semester Spanish students rather than second-semester students, as in Leow et al. (2008; see Participants section). Although this appeared to be the most appropriate level of administration, the scores from the control conditions from the two studies suggest that the level of participants may not have been completely comparable. The mean comprehension score for the control condition in Leow et al. was 46.5%, whereas the comprehension score for the control condition (of the TA group) in the current study was just 24%. Indeed, lower comprehension scores were found overall for the participants in this study. Given that independent measures of proficiency are not provided in this line of research, future research may consider piloting the control condition with an initial set of participants to determine whether participants at that level are likely to perform at a comparable level of comprehension as in previous studies.

The results of the study should also be considered in light of the fact that the TA group experienced a greater level of attrition than the NTA group, as reported in the Participants section. This was due to the fact that participants in the TA group were not as likely as participants in the NTA group to have circled at least 60% of the target forms, one of the criteria for inclusion in data analysis. It is necessary to note that this attrition occurred prior to the coding of the TAs and therefore is independent from the participant elimination based on noncompliance evidenced in TAs, which serves to raise the internal validity of a study (see Hama & Leow, 2010, note 2). Thus, although the internal validity of the study was not undermined by reactivity issues caused by TAs, the conclusions are restricted to the population represented by the participants who were able to circle at least 60% of the forms. Future studies that include a NTA group as a control for potential reactivity effects should also consider the rate of attrition (prior to participant elimination based on noncompliance evidenced in the TAs). Such systematic reporting of attrition is good practice for any empirical report (Leow, 1999).

A final general limitation of the current study is that its results are not directly comparable to results from VanPatten (1990), Greenslade et al. (1999), and Wong (2001), who all used the same passage (although

this passage was an English translation in Wong, 2001). As pointed out in the discussion, certain differences between the passages could have affected the results. Future research may address whether the length and possibly the difficulty of the passage itself has an effect on simultaneous attention to form and meaning in both the written and aural modes.

CONCLUSIONS

The present study examined both the theoretical issue of the role of attention in L2 acquisition and the methodological issue related to the reactivity of TAs as motivated by the line of research addressing the effect of simultaneous attention to form and meaning on L2 comprehension and by Leow et al. (2008) in particular. The results from the current study showed that thinking aloud led to a statistical but minimal effect that did not compromise the internal validity of the study. Additionally, the results indicated that, consistent with Leow et al. (2008) and Wong (2001), attending to grammatical or lexical form while reading a L2 text for meaning did not negatively affect comprehension, as might have been predicted by VanPatten's primacy of meaning principle. Indeed, on the basis of evidence from TA protocols, the findings suggest that L2 learners who processed these forms more deeply comprehended the passage. Future research in this area should consider additional factors, such as the timing of the input, L2 proficiency, and levels or depth of processing, so that researchers can arrive at a fuller understanding of the conditions under which attention to form is beneficial or detrimental to L2 comprehension as well as to L2 acquisition more generally.

(Received 9 February 2012)

NOTES

1. The current review of the literature focuses on the results related to the experimental versus the control conditions and not on comparisons between the experimental conditions themselves, as these comparisons are not directly related to the research questions of interest—that is, whether paying attention to form and meaning affects comprehension as compared to paying attention to meaning alone.

2. The corresponding author of Leow et al. (2008) provided the authors of the current study with the syllabus from the second semester of the Spanish language program in which the participants from that study were enrolled. On the basis of reported number of average hours of instruction, specifically 60 hr, it was determined that participants in Leow et al. had completed approximately two-thirds of the second semester and had likely completed 13 out of 17 *lecciones* “lessons” (approximately 76%) of their Spanish textbook, which was used over the course of two semesters. Participants in the current study had completed 42 of 53 *partes* “parts” (approximately 79%) of their textbook, which

was used over the course of three semesters. Both groups of participants had been instructed on the present tense, reflexives, object pronouns, forms and uses of the imperfect and preterit tenses, and informal commands. Participants in the current study had also been briefly introduced to aspects of the present and past subjunctive, formal commands, and the conditional tense.

3. Leow et al. (2008) report the passage to contain 358 words, but the researchers of the current study counted 354 words.

4. In the fall semester, an additional researcher was also involved in the data collection, scoring, and coding.

5. As an anonymous *SSLA* reviewer pointed out, translation of a form may reflect understanding of the form to a greater or lesser extent depending on the form (e.g., translation of the article *la* does not entail that the gender of the form was processed). The levels of processing adopted from Leow et al. (2008) are not posited to reflect particular cognitive processes (i.e., full understanding vs. partial understanding) but rather are used simply to “categorize levels of participants’ processes while interacting with the targeted forms in the input” (p. 679). See the Discussion section for a consideration of how future research may refine this approach.

6. Comprehension scores from participants from the fall and spring semesters were compared with a *t* test to ascertain that collecting data in two different semesters did not affect comprehension. The *t* test showed no significant differences between these groups of participants (fall semester: $n = 128$, $M = 0.351$, $SD = 0.153$; spring semester: $n = 223$, $M = 0.336$, $SD = 0.158$; $t(359) = .913$, $p = .362$).

REFERENCES

- Bowles, M. (2008). Task type and reactivity of verbal reports in SLA: A first look at a L2 task other than reading. *Studies in Second Language Acquisition*, *30*, 359–387.
- Bowles, M. (2010). *The think-aloud controversy in second language research*. London: Routledge.
- Bowles, M., & Leow, R. (2005). Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation. *Studies in Second Language Acquisition*, *27*, 415–440.
- Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, *62*, 73–101.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- 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.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.
- Ferguson, C. (2009). An effect size primer: A guide to clinicians and researchers. *Professional Psychology: Research and Practice*, *40*, 532–538.
- Greenslade, T. A., Bouden, L., & Sanz, C. (1999). Attending to form and content in processing L2 reading texts. *Spanish Applied Linguistics*, *3*, 65–90.
- Hama, M., & Leow, R. P. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition*, *32*, 465–491.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. London: Routledge.
- Leow, R. P. (1999). The role of attention in second/foreign language classroom research: Methodological issues. In J. Gutiérrez-Rexach & F. Martínez-Gil (Eds.), *Advances in Hispanic linguistics: Papers from the 2nd Hispanic Linguistics Symposium* (pp. 60–71). Somerville, MA: Cascadilla Press.
- Leow, R. P., Hsieh, H., & Moreno, N. (2008). Attention to form and meaning revisited. *Language Learning*, *58*, 665–695.
- Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, *26*, 35–57.

- McLaughlin, B. (1987). *Theories of second language learning*. London: Arnold.
- McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning*, 33, 135–157.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
- Polio, C., & Gass, S. (1997). Replication and reporting: A commentary. *Studies in Second Language Acquisition*, 19, 499–508.
- Polio, C., & Wang, J. (2005, October). *Another look at the reactivity of concurrent verbal protocols in second language reading research*. Paper presented at the Second Language Research Forum, New York.
- Rossomondo, A. E. (2007). The role of lexical temporal indicators and text interaction format in the incidental acquisition of the Spanish future tense. *Studies in Second Language Acquisition*, 29, 39–66.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
- Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Applied Language Learning*, 5, 57–93.
- Valdman, A. (1993). Replication study. *Studies in Second Language Acquisition*, 15, 505.
- Valentine, J. C., & Cooper, H. (2003). *Effect size substantive interpretation guidelines: Issues in the interpretation of effect sizes*. Washington, DC: What Works Clearinghouse.
- VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12, 287–301.
- VanPatten, B. (2004). Input processing in SLA. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–31). Mahwah, NJ: Erlbaum.
- VanPatten, B., Leiser, M. J., Keating, G. D., & Román-Mendoza, E. (Eds.). (2005). *Sol y viento: Beginning Spanish*. Boston: McGraw-Hill.
- Wong, W. (2001). Modality and attention to meaning and form in the input. *Studies in Second Language Acquisition*, 23, 345–368.

APPENDIX

Sample instructions. The following instructions were provided to participants who were asked to think aloud and were assigned to the *lo* attentional condition.

Please read the following text on the Aztecs for **comprehension**.

At the end you will be asked to answer some questions after your reading without referring back to the text.

As you read, please circle all instances of *lo*.

Example:

Antonio me lo dijo.

In addition, as you read the article and answer the questions, please **think your thoughts aloud**. That is, whatever passes through your mind while you **read the text for information and answer the questions**. You may speak either in English or Spanish.