Studies in Second Language Acquisition, 2010, **32**, 465–491.
doi:10.1017/S0272263110000045

LEARNING WITHOUT AWARENESS REVISITED

Extending Williams (2005)

Mika Hama and Ronald P. Leow
Georgetown University

The role of awareness or consciousness in learning has been a relatively contentious issue in non-SLA fields (e.g., cognitive psychology). With the publications of Williams (2004, 2005), a similar debate appears to be brewing in the field of SLA. Contrary to Leow (2000), who reported that unawareness did not appear to play an important role in second or foreign language development, Williams (2005) offered empirical evidence that learning without awareness appears to be feasible. At the same time, it is also noted that Leow's and Williams's (2005) research designs measured unawareness at different stages (online encoding and offline retrieval, respectively) of the acquisitional process. The present study revisited and extended Williams's (2005) study by using a hybrid design to gather concurrent data at the stage of encoding and during the testing phase as well as nonconcurrent data after the experimental exposure. Some methodological changes were also implemented to probe deeper into learners' thought processes. The quantitative analyses performed on the data of 34 carefully screened participants revealed that, at the encoding stage, unaware learners do not appear capable of selecting or producing the correct determiner-noun combination when required to do

We would like to thank John N. Williams for his critical review of this study, our statistician Rusan Chen for analyzing our data, and Melissa Baralt for her contribution in the taping of the experimental materials.

Address correspondence to: Mika Hama, Department of Linguistics, Georgetown University, 37th and O Streets, NW, Washington, DC 20057-1051; e-mail: mh278@georgetown.edu; or Ronald P. Leow, Department of Spanish & Portuguese, Georgetown University, ICC 403, 37th and O Streets, NW, Washington, DC 20057; e-mail: leowr@georgetown.edu.

© Cambridge University Press, 2010 0272-2631/10 $15.00

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

so from options that include both animacy and distance information. The qualitative data underscore the importance of not only situating the measurement of the construct (un)awareness from different stages—that is, both online and offline—but also triangulating data from several sources in any report on its role in learning. Plausible explanations for the differences in findings are discussed.

The role of awareness or consciousness in learning, generally subsumed under the rubric of implicit (without awareness) and explicit (with awareness) learning, is a relatively contentious issue in both non-SLA fields and SLA. For example, in cognitive psychology, there are two opposing views regarding the role of consciousness in learning. Some researchers claim that learning can occur without awareness (e.g., Curran & Keele, 1993; Litman & Reber, 2005; Nissen & Bullemer, 1987; Reber, 1976, 1993) or that learning without awareness may even be superior to learning with awareness (e.g., Reber, 1976, 1993). Others have questioned the claims about any disassociation between awareness and learning (e.g., Dulany, Carlson, & Dewey, 1984, 1985; Perruchet & Amorim, 1992; Perruchet & Pacteau, 1990, 1991; Perruchet & Pacton, 1996; Shanks, Green, & Kolodny, 1994; Shanks & St. John, 1994). With the publications of Williams (2004, 2005), a similar debate appears to be brewing in the SLA field. Contrary to Leow (2000), who reported that unawareness does not appear to play an important role in second or foreign language (L2) development, Williams (2005) offered empirical evidence that learning without awareness seems feasible. At the same time, it is also noted that Leow's and Williams's (2005) research designs measure unawareness at different stages of the acquisitional process: at the concurrent (online) stage of encoding the incoming information (Leow) and at the nonconcurrent (offline) stage of retrieval of stored knowledge of the construct (Litman & Reber)—that is, where learners indicate offline whether they were aware of the targeted underlying rule during the experimental exposure (Leow; Williams, 2005). The present study set out to investigate the role of unawareness at the concurrent stage of encoding the information provided during the training phase in addition to modifying some methodological aspects of Williams's (2005) research design to probe deeper into learners' processes. To this end, Williams's (2005) study was revisited by adding (a) a hybrid design to gather concurrent data at the stage of encoding and during the testing phase,[1] (b) items (four instead of two) on the multiple-choice (MC) test to extend learners' selection of options (animacy plus distance) in order to replicate the training context, (c) a production test, and (d) the same modality for both the training and testing phases.

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

AWARENESS AND LEARNING IN SLA

Similar to the cognitive psychology field, there has been no consensus on the role of awareness in L2 development in SLA, both at a theoretical level (e.g., Robinson, 1995; Schmidt, 1990, 1993, 1994, 1995, 2001; Tomlin & Villa, 1994) and at the level of empirical inquiry (e.g., Leow, 2000; Robinson; Rosa & Leow, 2004a; Williams, 2004, 2005). At the theoretical level, there are three major hypotheses on the association between awareness and L2 development in SLA. On one hand, Schmidt (e.g., 1990) argued that for input to be processed, a low level of awareness at the point of noticing some linguistic data in the input is necessary. In the noticing hypothesis, Schmidt stated that only data consciously attended to are processed and then become available for further processing. It should be noted that Schmidt posited two levels of awareness: awareness at the level of noticing—that is, "conscious registration of the concurrence of some event" (1995, p. 29)—and awareness at the level of understanding—that is, "recognition of a general principle, rule or pattern" (1995, p. 29), which is not required for input to be initially processed.

On the other hand, Tomlin and Villa (1994), in their functionally based, fine-grained analysis of attention in L2 processing from a cognitive science perspective, described attention as possessing three components with neurological correlates: alertness, orientation, and detection. Among these three components, they argued that detection, defined as the "cognitive registration of stimuli" (p. 190), is necessary for input to be processed and thus to become available for further processing. Crucially, according to their argument, awareness is not required for detection; in other words, it is not required for input to be available for language development.

Robinson (1995) reconciled these two opposing views of the role of awareness in L2 processing by postulating that although detection is important in input processing, it is at the level of noticing (i.e., attention plus minimally a low level of awareness) that linguistic data are available for further processing. To visually illustrate the differential roles of detection and noticing, Robinson hypothesized that detection occurs before noticing in the acquisitional process (see Leow, 2006, and Leow & Bowles, 2005, for comprehensive reviews).

EMPIRICAL STUDIES ON THE ROLE OF (UN)AWARENESS IN L2 DEVELOPMENT

Several empirical studies have examined the role of awareness in SLA (e.g., Leow, 1997; Robinson, 1995; Rosa & Leow, 2004a; Rosa & O'Neill,

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

1999; Sachs & Suh, 2007) and have generally reported beneficial effects of awareness on L2 development. Additionally, several levels of awareness have been postulated—namely, awareness at the level of noticing, at the level of reporting, and at the level of understanding (Leow, 1997; Rosa & Leow; Rosa & O'Neill; Sachs & Suh). To date, however, only a very limited number of SLA studies have directly examined whether language development can occur among unaware learners (Leow, 2000; Williams, 2004, 2005). Similar to the findings in cognitive psychology, the results of these studies reveal contradictory findings.

Leow (2000), who employed a hybrid design, reported that learning did not appear to occur among unaware learners. His study addressed the effect of awareness or lack thereof on L2 learners' subsequent intake and production of L2 forms. The targeted forms were third-person agreement of Spanish irregular verbs (ending in either *–er* or *–ir*) that have a stem change in the preterit. For example, the third-person singular of the preterit of *morir* "to die" does not follow the regular verb paradigm **morió* but *murió* "he died," where the stem vowel *o* changes to *u*; similarly, *pedir* "to ask for" is not **pedió* but *pidió* "he asked for," where the stem vowel *e* changes to *i*. Thirty-two English-speaking adults, learning Spanish as a foreign language at the college level, participated in a problem-solving task (crossword puzzle) in which they were exposed to 10 incomplete exemplars of the targeted forms interspersed among 15 clues. Prior knowledge (recognition and controlled written production) of the targeted forms was controlled for via a pretest administered 3 weeks before the task. The crossword puzzle was designed to require learners to fill in the correct endings of the irregular verbs without the need to pay attention to the stem change present in the incomplete targeted verb form (e.g., *mur–*).

Immediately upon completion of the crossword puzzle, participants engaged in a four-option MC recognition assessment task with items identical to those in the crossword clues and in a written production assessment task (a fill-in-the-blank) in which participants were to produce the targeted forms in different contexts (the same assessment tasks used for the pretest). To operationalize the construct of awareness, participants were instructed to think aloud nonmetalinguistically while completing the crossword puzzle and also during the postexposure assessment tasks. The verbal reports from the think-aloud protocols were then coded to establish whether a participant was either aware or unaware, based on the criteria posited in Leow (1997, 2001) for levels of awareness. Participants were assigned to the unaware group (their unawareness level was also cross-checked with offline awareness measures in the form of postexposure questions and of an interview) or to the aware group if they "provided a report of being aware of the target forms [a simple reference to the target forms that does not require mentioning of rules] or some form of metalinguistic description of the

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

underlying rule" (Leow, 2000, p. 564). The posttest scores of the aware participants' recognition and written production assessment tasks revealed significant gains from the pretest scores on these tasks. In contrast, among unaware learners, there were no significant differences between their performances on the pretest and posttest. Leow concluded that "no disassociation between awareness and learning was found in this study" (2000, p. 573).

Williams (2004, 2005) reported that unaware learners can provide evidence of learning without awareness. Improving the methodology used in his 2004 study, Williams (2005) conducted two experiments to test whether participants were able to learn miniature noun class systems without awareness. The participants in Williams's (2005) experiments were from a variety of language and linguistics-related backgrounds. The noun phrases used in the experiments were four novel determiners: *gi*, *ro*, *ul*, *ne*. *Gi* and *ro* are the English translation equivalence of "near" and *ul* and *ne* are equivalent to "far." These determiners also carry animacy values: *Gi* and *ul* are animate and *ro* and *ne* are inanimate.

Before the training phase, participants were provided with the translation of the four novel determiners for "near" or "far"; however, they were not informed about the animacy rules of these determiners. During the training phase, the participants were presented with individual sentences with noun phrases that comprised a novel determiner and an English noun (e.g., ***gi*** *dog* "near dog"). For each sentence, they (a) listened to the sentence, (b) indicated if the novel word meant "near" or "far," (c) repeated the sentence aloud, and (d) were requested to create a mental image of the situation described by the sentence. The participants were told that the main focus of the study was memory to distract them from paying attention to the underlying animacy rules.

After the training phase, to test participants' learning of the animacy rules, participants engaged in a written selection assessment task with new sentences that contained either trained or untrained (new) noun phrases. These sentences were incomplete because the targeted structure (i.e., the noun phrase) was missing. The participants were then given two noun phrases (one animate and one inanimate) and instructed to select in a two-option MC assessment task the noun phrase that seemed "'more familiar, better, or more appropriate' on the basis of what they had heard during the training task" (Williams, 2005, pp. 282–283). Upon completion of the MC assessment task, the participants were asked what criteria they had used to make these choices. If they did not mention "any references to living or nonliving, moves or does-not-move, and so forth" (p. 283) as a reason for their choice, they were classified as unaware. The results revealed that many of the participants who were classified as unaware were able to choose the correct noun phrase during the assessment task at a level significantly above chance.

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

The conflicting results reported in the two studies are not surprising, given the many differences between Leow's (2000) and Williams's (2005) studies. The participants are from different language backgrounds and exposed to different targeted items and experimental learning tasks. To measure learning, Leow employed both a four-option MC recognition and a written production assessment task, whereas Williams used a two-option MC assessment task. Awareness in Williams's study was coded at the level of understanding, whereas Leow's study was at a lower level—namely, that of noticing.

However, the primary differences between Williams's (2005) and Leow's (2000) studies lie in the methodology used to measure the construct unawareness and the stage at which unawareness was measured. Leow used a hybrid design to measure unawareness both at the encoding stage with an online procedure (verbal reports) and at the retrieval stage with an offline questionnaire. Williams's measurement of unawareness was at the retrieval stage via an offline procedure (questionnaire), which asked learners to verbalize rules after the completion of the training and testing phases. The use of verbal reports, especially metalinguistic ones, has been critiqued for its potential impact on the alteration of learners' internal processes because of the verbalization (the issue of reactivity; Ellis, 2001). In other words, the act of thinking aloud during task performance may alter participants' internal processes. Current studies on the issue of reactivity have provided some conflicting findings in the field of SLA: Four have reported nonreactive findings (Bowles, 2008; Bowles & Leow, 2005; Leow & Morgan-Short, 2004; Sachs & Suh, 2007), one has reported negative reactivity, although this study differs because the L2 was used in the verbal report (Sachs & Polio, 2007), and one has reported positive reactivity (Sanz, Lin, Lado, Bowden, & Stafford, 2009).

Offline data elicitation procedures have been criticized by cognitive psychology researchers as being an insensitive measure of awareness (Shanks et al., 1994; Shanks & St. John, 1994). For example, Shanks et al. found that participants who were classified as unaware by means of offline verbal reports performed significantly better on tests compared to unaware learners who were classified by another type of awareness measure—the cued recall, which is also frequently used in cognitive psychology studies of awareness. This validity issue may be a result of memory decay and fabrication, or both, which are claimed by many researchers to be major deficiencies in offline measurements (Ericsson & Simon, 1984, 1993; Leow, 2000; Leow & Morgan-Short, 2004). According to Ericsson and Simon (1993), learners may not be able to access the information stored in short-term memory due to the time lag between the event and the recall when using offline protocols; therefore, memory decay may affect their postexposure verbal report. In contrast, using online verbal protocols to address unawareness at the encoding and concurrent stage is likely to be valid because the information is still fresh

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

in working memory and directly accessible for verbal reports. Additionally, SLA researchers, such as Bialystok (1979) and Ellis (2004), cautioned against the assumption that rule verbalization is an accurate measure of awareness. Therefore, the research methodology and the stage at which the construct of awareness is measured are of paramount importance, given the obvious difference between the type of data elicited at each stage: One is concurrent, whereas the other is nonconcurrent.

To extend Williams's (2005) study to the concurrent stage of encoding, the present study used a hybrid design that included both online and offline data elicitation procedures to gather data at the concurrent stage of encoding, during and after the testing phase. Additionally, Williams's research design was extended to probe deeper into learners' processes by (a) increasing the number of items (four instead of two) on his MC test to include the presence of distance in learners' selection of options (i.e., animacy plus distance) in order to replicate the training context, (b) including a production test in addition to the MC test, and (c) providing the same modality for both the learning and testing phases. This study was guided by the same research question as Williams's study: Is learning without awareness (at the encoding stage) possible?

METHODOLOGICAL CHANGES TO THE RESEARCH DESIGN

The Use of Concurrent Data Elicitation Procedures (Think-Aloud Protocols)

Studies that examine the concurrent stage of encoding have employed think-aloud protocols in research designs to operationalize and measure awareness in SLA. Think-aloud protocols are verbal reports elicited from participants while they are interacting with the L2 input. These protocols can be of several types that include nonmetalinguistic protocols (i.e., participants focus on the task at hand and only voice their thoughts without providing any explanation for them) or metalinguistic or metacognitive protocols (i.e., participants also provide some additional information such as reasoning or explanation or a metacognitive report on what their thought processes are). Empirical evidence supports the value of operationalizing attention and awareness using concurrent measures (e.g., online think-alouds in Bowles, 2003, 2004, 2008; Leow, 1997, 1998a, 1998b, 2000, 2001; Leow & Morgan-Short, 2004; Rosa & Leow, 2004a, 2004b; Rosa & O'Neill, 1999; Rott, 2007; Sachs & Suh, 2007). The present study included online measures (think-aloud protocols) that would help to establish any reported awareness from the data at the concurrent stage of encoding in addition to providing important concurrent data that would shed light on how learners were processing the input during the training and testing phases.

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

The MC Assessment Task

In Williams's (2005) MC assessment task, the participants were instructed to choose between two options (two items that differed in terms of animacy: e.g., ***gi*** [animate-near] *cushions* or ***ro*** [non-animate-near] *cushions*). Williams reported the elimination of several participants who demonstrated awareness of the targeted rule during the testing phase. The present study included an additional two choices for the selection assessment task by incorporating the distance dimension (e.g., ***gi*** [animate-near] *cushions*, ***ro*** [non-animate-near] *cushions*, ***ul*** [animate-far] *cushions*, and ***ne*** [non-animate-far] *cushions*). This change was implemented not only to maintain a context comparable to the original training phase to probe deeper into learners' processes but also to avoid the test effect of learners potentially becoming aware of the animacy rule during this testing phase, as reported in Williams's study.

The Production Task

The rationale for the use of multiassessment tasks to measure the process of learning has been well documented in the SLA literature (e.g., Leow, 1998a; Shook, 1994; VanPatten & Sanz, 1995). The present study included a production assessment task in addition to the MC assessment task to measure participants' learning of the animacy rule.

Modality

The test sentences in Williams's (2005) MC assessment task were provided in the written mode, whereas the sentences in the training task were provided orally. For the present study, the modality of presentation between the training and testing phases remained constant; that is, sentences were provided orally for both the training and testing tasks.

METHOD

Participants

To ensure that participants were representative of the experimental condition of the study—namely, that they had learned the distance rule

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

and were reportedly unaware of the animacy rule—several participants were excluded for various reasons. These participants (a) failed to complete or inaccurately completed the training or assessment tasks (n = 34); (b) failed the prelearning vocabulary test (n = 8); (c) indicated the animacy rule, either online or offline (n = 9); or (d) clearly demonstrated a non-animacy-based strategy that would result in obtaining 50% accuracy on the critical items on the assessment tasks (n = 11). This screening led to the selection of 34 participants.[2]

Participants completed a language background questionnaire that asked for their age, first language (L1), L2 experiences, and linguistics training. This information was used to eliminate participants who had linguistics backgrounds.

Materials

Noun Phrases. Due to the inclusion of a production assessment task, eight nouns were added to the noun phrases used in Williams's (2005) study. Thus, 16 new noun phrases were created for the present study (see Table 1). These new noun phrases were used for the training task (n = 8) and the production task (n = 8).

The additional nouns were carefully chosen according to the criteria described in Williams's (2005) study: Excluding "natural things that do not typically move (e.g., trees or rocks) and man-made things that do move (e.g., bicycles or cars)" (Williams, p. 280). These criteria are important to prevent participants from creating other types of categorizations of these nouns (e.g., natural vs. man-made, moves vs. does-not-move).

For each class, there were 16 nouns. During the training task, eight nouns of each class occurred with both possible determiners (*gi*, *ro*, *ul*, *ne*) in either singular or plural form. The other eight nouns of each class only occurred with one determiner during the training task (critical items). The combinations that were not used in the training task were used for the assessment tasks to assess participants' generalization

Table 1. Additional noun phrases used in the present study

Living		Nonliving	
Near	Far	Near	Far
gi turtle(s)	*ul* turtle(s)	*ro* lamp(s)	*ne* lamp(s)
gi horse(s)	*ul* horse(s)	*ro* spoon(s)	*ne* spoon(s)
gi rabbit(s)	*ul* rabbit(s)	*ro* candle(s)	*ne* candle(s)
gi elephant(s)	*ul* elephant(s)	*ro* phone(s)	*ne* phone(s)

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

abilities (new test items). To ensure that all nouns appeared an equal amount of times, the critical items occurred in both singular and plural forms. Singular and plural nouns in each class appeared with equal frequency in both the training task and the assessment tasks.

Sentences. The sentences used during the training task were also adapted from Williams's (2005) study with some modifications. New sentences were created for the eight new noun phrases added to the training task (see Appendix A). The sentences in the training task were divided into two sets. Eight additional new sentences were created for the eight new noun phrases, which were generated for the new assessment task (see Appendix B).

Procedure

The procedure used here generally followed that of Williams (2005). With the exception of the language background questionnaire and the postexposure questionnaire, all tasks were completed on a randomly assigned computer. The verbal reports from the think-alouds were recorded digitally on the computer and the postexposure questions were recorded on paper.

Vocabulary Pretraining Task. After being presented with the four novel determiners, the participants engaged in a two-step vocabulary pretraining phase. First, the participants practiced the novel words until they were able to verbalize the word when presented with its English translation. Second, they listened to one of the novel words recorded by a native English speaker and indicated the English translation of the word, by selecting one of the two options, "near" or "far," presented on a computer screen. Each word was presented 12 times in random order. In addition to Williams's (2005) pretraining design, during the vocabulary pretraining task, learners were instructed to think aloud while they made choices between "near" and "far." The pretraining vocabulary test was given to ensure that participants in fact learned the four novel words.

Training Task. Upon completion of the vocabulary pretraining task, participants received these instructions adapted from Williams (2005):

> This is an experiment about how people formulate sentences in different languages. Different languages have different ways of saying things, which raises the question of whether this forces people to think slightly differently in order to speak and understand in different languages. Here I am investigating this issue in a situation where the sentences are almost

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

> entirely in English, apart from the four words that you have just learned. (pp. 281–282)

An example sentence, *The little boy patted **gi** tiger in the zoo*, not used in either the training or assessment tasks was presented to the participants along with additional instructions:

> In this language, each time an object is mentioned it is necessary to specify how far away it is from the subject of the sentence, whether it is "near" or "far." The word that is used to do this also functions like the English "the." So saying *gi tiger* is like saying "the-near tiger." In the context of this example, the word expresses the idea that the boy is near the tiger (obviously, because he is patting it). (Williams, 2005, p. 282)

After the series of instructions, participants were informed that they would see many different sentences and that their memory of these sentences would be tested at the end. However, participants were also informed that memorization of the exact wording of these sentences was not necessary for successful completion of the memory test and that the understanding of the general situations described by the sentences was sufficient. Then, participants were asked to complete the following steps: (a) listen to a sentence, (b) repeat it aloud exactly, (c) choose "near" or "far" by selecting one of the choices while providing verbal reasoning for their choices by thinking aloud, and (d) form a mental image of the general situation described in the sentence. The sentences in the training task were presented in the following order: the sentences from the first set followed by the sentences from the second set. This presentation was repeated two more times. The sentences in each set were independently randomized for each participant.

Assessment Tasks

MC assessment task. The sentences from Williams's (2005) revised test sentences, which were revised during his experiment, were used for this task with one modification. This modification was necessary due to the increase in the number of choices for the MC assessment task in the present study. In Williams's study, some of the sentences in the MC assessment task did not provide distance information for the object (e.g., *The wolf wanted to eat **ul** birds*). Because the selection task used in his study only required the participants to make a choice between two items that were different only in terms of animacy (e.g., ***ul*** [animate-far] *birds* or ***ne*** [inanimate-far] *birds*), this lack of distance information did not create any problem. To extend the option of an either-or response based on the choice of animacy, Williams's sentences were modified to allow the context to infer the distance of the object (e.g., *The wolf wanted to eat **ul** birds on top of the tree*). More

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

specifically, the present study's MC assessment task increased the number of choices from two to four by incorporating the dimension of distance—for example, ***gi*** [animate-near] *birds*, ***ro*** [inanimate near] *birds*, ***ul*** [animate-far] *birds*, and ***ne*** [inanimate-far] *birds*. To be able to choose an option from the four choices, the participants must receive information on the distance of the object in relation to the sentence (see Appendix B).

In the present study, the MC assessment task contained eight new noun phrases not used in the training task and eight noun phrases previously used in the training task. Two phrases from the training task were used as training items and were presented at the beginning of the task. The eight new items were presented first, followed by the eight trained items, and then followed by another eight new items. Due to the change in the sentences used for this task (i.e., incorporating distance information in these sentences), the testing procedure was modified in the following manner. The participants in Williams's (2005) study received an incomplete sentence (e.g., *I had to read the manual to find out how to adjust…*) and were then given two alternative completions (e.g., ***gi*** *clock* or ***ro*** *clock*) from which to choose. Here, because the sentences used incorporated distance information, it was not possible to complete the task in the same manner. Thus, for each item, participants listened to the sentence in which the object was missing its determiner (a short beep sound indicated the position of the missing determiner). Participants were then asked to select the correct determiner from among the four options (*gi*, *ro*, *ul*, *ne*) while they thought aloud.

(1) An MC test item
I had to read the manual to find out how to adjust [beep] *clock up on the wall.*
Choices: (a) ***gi*** *clock*, (b) ***ro*** *clock*, (c) ***ul*** *clock*, (d) ***ne*** *clock*

Production assessment task. New sentences were created for the eight new noun phrases, which were generated for this new assessment task. These new sentences were designed to provide distance information and were created to ensure that the general situation depicted by the sentences used in the training task and those used in this production task were as dissimilar as possible.

The production task also contained eight noun phrases not used in the training task (they were also different from those on the MC assessment task) and eight noun phrases previously used in the training task. Two sentences from the training task were presented before the new and trained items, in the same order as in the MC assessment task. The eight new items were presented first, followed by the eight trained items, and then followed by another eight new items. For each item (both the trained and new), participants listened to a sentence in which

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

the object was missing a determiner (a short beep sound indicated the position of the missing determiner). Participants were then asked to provide the missing determiner and were instructed to think aloud during this task.

(2) A production test item
In the gallery we all admired [beep] *pictures from the other side of the room.*

To avoid additional exposure to the targeted items, the production assessment task preceded the MC task.

Awareness Measures

Concurrent Versus Nonconcurrent. The concurrent awareness measures came from verbal reports from the think-aloud protocols (online measure). The think-alouds were considered the main measure of unawareness because they provide accessibility to the construct of awareness at the stage of encoding.

The nonconcurrent awareness measure was similar to Williams's (2005) offline questionnaire and was administered upon the completion of the assessment tasks. Participants were asked to disclose the specific criteria they had used to make their choices on the assessment tasks. The postexposure reports were also used to ensure that participants who did not produce any verbal report did not in fact have knowledge of the targeted animacy rule.

Coding of Awareness. The recorded verbal reports from the think-aloud protocols during the training and assessment tasks were transcribed and coded as *understanding*, *noticing*, or *no report* based on the coding scheme in Rosa and O'Neill (1999). A verbal report was coded as (a) noticing when some aspect of animacy was mentioned or commented upon, (b) understanding when correct rules related to animacy were mentioned, or (c) no report when the report did not fall under the coding categories of noticing or understanding. Interrater reliability between two raters was 100%. The reports from the questionnaire were also coded for awareness along similar lines.

Learning Measurements

Participants' learning was measured by their performances on the MC and the production assessment tasks. Following Williams (2005),

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

a correct answer for a test item on both the MC and the production tasks was awarded one point; an incorrect answer received no points.

RESULTS

Qualitative Analyses

Concurrent Data. During the training phase, there was no reported instance of awareness of the animacy rule. The think-aloud protocols clearly established that participants were using a distance bias in their selection of choices during this phase. A typical report is provided in (3).

(3) *The old woman liked to sit for hours with **gi** cat on her lap* (repeating the sentence heard). *Near because it's on her lap.*

During the testing phase, at the level of noticing, the think-aloud protocols revealed five reported instances of participants who made a comment on some aspect of animacy. These participants provided either a partial report or an incorrect rule of assigning animacy to the novel words, as illustrated in (4).

(4) a. ***Ne** birds because **ne** represents living thing.*
b. ***Ul** table, it's nonliving thing, that's far away.*
c. ***Ne** because it's an animal and far away.*
d. ***Ul** book no… **ne** book. **Ne** and **gi** indicate you know **ne** and **ul**. I think one indicates animal and one indicates nonliving thing, I think **ne** represents ah nonliving things.*

At the level of understanding, there were three instances of participants, provided in (5), who made references to the correct underlying animacy rules as they performed the testing task.

(5) a. ***Gi** is referring to animals, I think.*
b. *I think **ul**s are frequently used with far animals in previous examples, **gi** also with animals but when they are close because animal and close near mouse **gi.***
c. ***Ne** cushions because it's inanimate, I think.*

This level of understanding was corroborated on the exit questionnaire, as illustrated in (6).

(6) a. ***Gi**-near animal, **ro**-near object, **ul**-far animal, **ne**-far object.*
b. *Near animal = **gi**, Far animal = **ul**, Near things = **ro**, Far things = **ne**.*

Of the remaining five participants (at the level of noticing), three also reported some reference to the animacy rule (whether correct or

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

incorrect) on the offline questionnaire, whereas two did not, as shown in (7).

(7) *The thing that I did is that I associated the first two,* ***ne*** *&* ***ul*** *with far away and* ***gi*** *&* ***ro*** *with near and I had that ingrained in my mind.*

Additionally, another participant reported awareness of the animacy rule at the level of noticing on the postexposure questionnaire but not during the think-aloud protocol. As in Williams's (2005) study, participants who reported awareness at the level of understanding demonstrated a high level of accuracy on the MC (79% and 81% for the trained and new items, respectively) and production (92% and 96% for the trained and new items, respectively) tasks. These nine aware participants were subsequently removed from the participant pool.

Non-Animacy-Based Strategy. The think-aloud protocols collected during the testing phase also revealed that two participants clearly relied on a non-animacy-based strategy to select the expected option. Here, strategy is defined as a conscious effort on the part of the learner to learn something, which arguably would place a participant in the aware group. An example of such a strategy is provided in (8).

(8) a. *I have decided that* ***ne*** *is always far,* ***ro*** *is always near. The other two will be excluded for obvious reasons.*
b. *Near* ***ro*** *I'm just sticking to* ***ro*** *for near. I hope that's OK.*

Similar to what the think-aloud protocols gathered on awareness during the testing phase, the corresponding response on the offline questionnaire of this participant, for example, does not fully reflect her thought processes during this phase.

(9) *Color of word.* ***Ne*** *= far (sounds alike). Imagining the situation from before.*

The exit questionnaire also produced three exemplars of non-animacy-based strategy use, as illustrated in (10).

(10) a. *I used the criteria of the two fars and the two nears as singular and plural. I did not pick up on this until near the end of the study. For example,* ***gi*** *snakes: I used the* ***gi*** *forms as the plural one.*
b. *I thought the* ***ro*** *and* ***ul*** *were plural and that* ***gi*** *and* ***ne*** *were singular and that* ***ro*** *and* ***gi*** *were near and* ***ul*** *and* ***ne*** *were far.*

Based on these data, a careful examination of the assessment tasks revealed six additional instances of non-animacy-based strategy use. These 11 participants were subsequently removed from the participant pool.

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

Quantitative Analyses

MC Assessment Task. To establish whether unaware participants were able to select the correct determiner-noun combination of trained items at a level significantly above chance on the four-option MC assessment task, a one-sample *t* test was conducted, which revealed that this group performed significantly better than chance on the MC assessment task, $t(33) = 6.475$, $p < .001$. The scores obtained on the MC assessment tasks were then submitted separately for trained and new items to 2 × 2 repeated-measures ANOVAs with animacy (correct or incorrect) and distance (correct or incorrect) as within-subject factors. The dependent variables were types of responses—namely, correct animacy and correct distance (CACD), correct animacy and incorrect distance (CAID), incorrect animacy and correct distance (IACD), and incorrect animacy and incorrect distance (IAID). All analyses were run using the general linear model program on the statistical package for social sciences (SPSS). Mean scores and standard deviations for the MC assessment tasks for both trained and new items are reported in Table 2, and the ANOVA results for each type of items are provided in Table 3.[3]

Table 3 shows that there were main effects for distance for both trained, $F(1, 34) = 187.17$, $p < .0001$, and new, $F(1, 34) = 380.76$, $p < .0001$, items with large effect sizes of .85 and .92, respectively. No main effects were found for animacy for either trained, $F(1, 34) = 0.801$, $p = .377$, or new, $F(1, 34) = 0.646$, $p = .427$, items nor were there significant interactions between distance and animacy for either trained, $F(1, 34) = 2.30$, $p = .138$, or new, $F(1, 34) = 1.95$, $p = .172$, items. The critical distinctions between type of responses to address whether learning took place without awareness are between the means of responses of the CACD and the IACD responses for both the trained (3.62 vs. 3.03) and new (6.94 vs. 7.56) items, respectively. Paired-samples *t* tests revealed no significant differences between these two types of responses, $t(33) = 1.340$, $p = .189$, and $t(33) = -1.187$, $p = .244$, respectively, which indicates that distance

Table 2. Mean items correct and standard deviations of the trained ($N = 8$) and the new ($N = 16$) items by type of response

	Trained items		New items	
Type of response	*M*	*SD*	*M*	*SD*
CACD	3.62	1.457	6.94	1.757
CAID	0.62	0.922	0.82	1.359
IACD	3.03	1.359	7.56	1.845
IAID	0.71	1.001	0.68	0.945

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

Table 3. ANOVA for the trained and new items

	Trained items			New items		
Source	*F*	*p*	η^2	*F*	*p*	η^2
Animacy	0.722	.377	.021	0.646	.427	.019
Distance	183.33	.000*	.847	380.759	.000*	.920
Animacy × Distance	2.499	.123	.070	1.952	.172	.056

*$p < .0001$.

was the overriding factor in participants' selection of responses. In other words, provided with the distance context in addition to the animacy option, unaware participants did not appear to demonstrate significant learning of the animacy feature of the novel words at the selection level of trained items nor at the level of generalization to new items. Paired-samples *t* tests conducted on the following trained versus new items comparisons for the CACD and IACD responses, respectively, revealed one significant result: Participants selected significantly more IACD responses for new items when compared to trained items ($p < .05$), which indicates that participants rely significantly more on the distance bias during the generalization task.

Production Assessment Task. To establish whether unaware participants were able to produce (out of four options) the correct determiner-noun combination of trained items at a level significantly above chance on the production assessment task, a one-sample *t* test was conducted, which revealed that this group performed significantly better than chance on the production assessment task, $t(33) = 8.023$, $p < .001$. Mean scores and standard deviations for the production assessment tasks for both trained and new items are provided in Table 4, and the ANOVA results for each type of items are reported in Table 5.

Table 4. Mean items correct and standard deviations of the trained ($N = 8$) and the new ($N = 16$) items by type of response

	Trained items		New items	
Type of response	*M*	*SD*	*M*	*SD*
CACD	4.06	1.496	7.47	2.596
CAID	0.35	0.646	0.59	0.892
IACD	3.24	1.458	7.18	2.713
IAID	0.35	0.646	0.76	1.182

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

Table 5. ANOVA for the trained and new items

	Trained items			New items		
Source	*F*	*p*	η^2	*F*	*p*	η^2
Animacy	2.962	.095	.082	0.019	.890	.001
Distance	282.754	.000*	.895	443.555	.000*	.931
Animacy × Distance	2.962	.095	.082	0.275	.604	.008

*$p < .0001$.

As illustrated in Table 5, the data on the production assessment tasks yielded results similar to the ones on the MC assessment tasks. There were main effects for distance for both trained, $F(1, 34) = 282.754$, $p < .0001$, and new, $F(1, 34) = 443.555$, $p < .0001$, items with large effect sizes of .89 and .93, respectively. No main effects were found for animacy for either trained, $F(1, 34) = 2.962$, $p = .095$, or new, $F(1, 34) = 0.019$, $p = .890$, items, nor were there significant interactions between distance and animacy for either trained, $F(1, 34) = 2.962$, $p = .095$, or new, $F(1, 34) = 0.275$, $p = .604$, items.

The critical distinctions between type of responses to address whether learning took place without awareness are also found between the means of the CACD and the IACD responses for both the trained (4.06 vs. 3.24) and new (7.47 vs. 7.18) items. Paired-samples *t* tests revealed no significant differences between these two categories, which indicates that distance was the overriding factor in participants' production of responses. In other words, provided with the distance context in addition to the animacy option, unaware participants did not appear to demonstrate significant learning of the animacy feature of the novel words at the production level of trained items or at the level of generalization to new items.[4]

DISCUSSION

Unaware learners, at the stage of encoding, did not appear to demonstrate any significant animacy bias in either the selection or the production of the trained or new determiner-noun combinations. The findings demonstrated instead a significant distance bias that reflected the overt focus of the training phase. Note that on the generalization tasks (new items), the participants demonstrated a significant increase on the dependency on distance (i.e., the IACD response) for their selection of the determiner-noun combination when compared to the trained items. In

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

sum, provided with options that included both animacy and distance, these participants' selection and production of the correct determiner-noun combination for both trained and new items appeared to be overridden by the choice of distance over animacy, which replicates the condition under which they learned the novel words. Overall, these results corroborate those reported in Leow (2000), who found no disassociation between awareness and learning.

Williams (2005) found that "unaware participants were able to select the correct determiner-noun combination at significantly above-chance levels" (p. 287) at the retrieval stage when participants were asked to select one of two options that featured only the animacy distinction. Several explanations for the difference between Williams's findings and those of the current study will be discussed in light of (a) the research methodology used, (b) the number of options from which to select in the MC assessment task, (c) the increase of items in the training phase, and (d) the modality of the training and testing phases.

Recall that Williams (2005) measured unawareness at the nonconcurrent stage of retrieval with an offline questionnaire; that is, data were gathered on the availability of learners' stored knowledge of the construct awareness after exposure to the training phase. Concurrent measurement of unawareness provides data on the accessibility of information regarding learner awareness at the encoding stage. In this study, a hybrid design was used to carefully screen the participant pool with both online (think-aloud protocols) and offline (exit questionnaire) measures to ensure that they relied on a distance bias in their selection of choices during the training phase and did not report awareness of the targeted rule during the testing phase. The concurrent data gathered during the training phase revealed no reported instance of awareness, which provides empirical online evidence to support Williams's proposal that, with the exception of a few participants, the learners in his study were also unaware of the animacy rule during the training phase. This finding may not be surprising given that participants were asked to focus solely on the distance information during this phase, as evidenced in the think-aloud protocols.

However, although both the offline questionnaire data and the verbal reports gathered while participants were completing the assessment tasks clearly revealed instances of learner awareness at both the levels of noticing (six instances) and of understanding (three instances), such important information may not be found in offline measures alone. For example, the concurrent data of one participant gathered during the testing phase revealed that the construct of awareness may be captured during a think-aloud protocol but not reported on an offline measure as exemplified in the following protocol reported in Table 6. As illustrated in Table 6, this participant began the task with a clear focus on the distance information, began to report awareness of the animacy rule for

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

Table 6. Think-aloud protocol

Correct answer	Actual answer	Correct animacy?	Correct distance?	Think-aloud comments
Ul *dog*	*Ne*	No	Yes	*Ne dog* because it's far away
Gi *bears*	*Ro*	No	Yes	*Ro bears* because it's very close
Ro *television*	*Ul*	No	No	*Ul television* because it's… yeah *ul* works
Gi *mouse*	*Gi*	Yes	Yes	The cat caught the *gi mouse*. Because it's close. And *gi* is referring to animals, I think.
Ne *sofa*	*Ne*	Yes	Yes	The *ne sofa* because it's far away
Ul *flies*	*Ne*	No	Yes	The chameleon uses (inaudible) *ne*. That means far far. *Ne*
Ne *cushions*	*Ne*	Yes	Yes	*Ne cushions* because *it's inanimate I think*
Ro *plates*	*Ro*	Yes	Yes	*Ro plates* because they are close
Gi *rabbits*	*Gi*	Yes	Yes	The *gi rabbit* because it's close and an animate object
Ne *spoons*	*Ne*	Yes	Yes	Ah *ne spoons* because it's far away
Ul *turtle*	*Ul*	Yes	Yes	*Ul turtle* because it's far away and yeah it's an animal
Ro *phone*	*Ro*	Yes	Yes	That would be it would be *ro phone* because it's near
Ul *horses*	*Ne*	No	Yes	*Ne horses* because they are far away
Gi *elephants*	*Ro*	No	Yes	*Ro elephants* because they are close to you
Ne *lamp*	*Ul*	No	Yes	*Ul* because it's ah ah next room
Ro *candles*	*Ro*	No	Yes	*Ro candles* because they are close to you

three of the four novel words (*gi, ul, ne*), and then ended up by reverting back to the distance information once more for his selections. On the exit questionnaire administered immediately after the assessment task, he reported on the distance bias only: "The thing that I did is that I associated the first two, ne & ul with far away and gi & ro with near and I had that ingrained in my mind," a report that would classify him as unaware.

Additionally, the presence of a non-animacy-based strategy while performing these tasks was reported both in the concurrent data (two instances) and nonconcurrent data (three instances). A careful review of the data revealed an additional six instances of a non-animacy-based strategy and that 8 of 11 of the participants who used this strategy failed to report its use on the questionnaire. The potential to rely on a nonreported strategy instead of an animacy-based rationale cannot be ruled out. The reported mismatches between concurrent and nonconcurrent data appear to support the call of previous literature not only for finergrained measurements of constructs such as awareness or lack thereof (Ericsson & Simon, 1984, 1993; Leow, 2000; Leow & Morgan-Short, 2004; Shanks et al., 1994; Shanks & St. John, 1994; Williams, 2005) but also for the important need for triangulating data from several sources in any report on the role of awareness in learning (Leow).

From a methodological perspective, the use of a hybrid design to address the issue of unawareness contributed substantially to (a) the operationalization and measurement of unawareness at both stages of encoding and retrieval, (b) the careful screening of participants who fulfilled the criterion of being unaware, and (c) the provision of deeper insights into learners' processes as they interacted with the input in both the training and testing phases. Measuring the construct of unawareness at both stages revealed that there may be mismatches between what learners became aware of while processing the input and the information they provided offline.

The provision of two options versus four on the MC assessment task also provides another plausible explanation for the difference in results. In Williams's (2005) assessment task, participants only had to choose between an animate or inanimate option with the feature distance not included in these options. At the stage of retrieval, Williams reported learning without awareness. The present study extended Williams's design to probe whether unaware participants would also be able to demonstrate the ability to select or produce the appropriate determiner-noun combination within the context of distance, as presented in the training phase. To address this issue, two options of distance were added to the original two options of animacy, which made the MC test a four-option task.

One clear difference between two versus four options may have to do with level of cognitive effort required to process the input. Whereas

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

learners in Williams's (2005) study only needed to pay attention to a noun phrase distinction of animacy with a 50% chance, learners in the present study needed to process the incoming information not only at a sentential level to accommodate both animacy and distance information in the assessment prompt but also to select or produce one out of four options (25% chance). Similarly, given that the training phase was longer than the original study with the inclusion of 16 new noun phrases, it may be argued that this phase was more cognitively challenging when compared to the original design. However, when the different number of options is addressed statistically, learners' performance or memory of trained items in this study is relatively similar to that of learners in the original study (69% vs. 70% on the MC task, respectively). Therefore, the increase in items during the training phase may not have contributed significantly to the difference in results.

Modality of presentation of the training and testing phases may also have played a role in the differential results of the two studies. Recall that the participants in Williams's (2005) study were exposed to the targeted novel combinations in the aural and written modes in the training and testing phases, respectively, whereas both the training and testing phases were presented orally in the present study. Previous literature on the role of modality and intake (e.g., Leow, 1995; Wong, 2003) has suggested potential differences in the way learners process the information received depending on the mode in which it is presented. Listening to the assessment prompts could have potentially been a more challenging task than a written format, which provides an additional visual exposure to the targeted items and which could have allowed learners to quickly reread the prompts. Therefore, the role of modality as a contributing factor to the differential results cannot be overlooked.

LIMITATION AND FUTURE RESEARCH

One limitation was the use of only two postexposure assessment tasks that were relatively close in format, as evident in participants' similar performances on both tasks. Future studies may need to include a wider range of assessment tasks (e.g., beyond the sentential level) to further address learners' differential behavior on tasks. Future research may also want to include a non-think-aloud group to address the potential effect of reactivity, although the similar behavioral performances on the trained items between the participants in Williams's (2005) and in the present studies do not appear to strongly support this research direction. The concept behind Williams's design clearly warrants future investigation into the role of unawareness in learning beyond the sentential level.

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

CONCLUSION

The present study was conducted to revisit the role of unawareness in learning specifically in relation to Williams's (2005) well-conceived study on this issue, given the inconclusive findings in this strand of SLA research. More specifically, by adding the use of online data elicitation procedures to Williams's original research design, this study addressed unawareness at the concurrent stage of encoding to probe more deeply into the issue of its role in learning. In addition to some methodological modifications, the think-aloud protocols were used to gather concurrent data to investigate learners' processes both during the training and the testing phases.

The quantitative analyses revealed that, at the encoding stage, unaware learners do not appear capable of selecting or producing the correct determiner-noun combination when both animacy and distance information was included. This inability of unaware learners to select or produce the correct determiner-noun combination in such instances may be due to the type of research methodology used here, to the level of cognitive effort required to process the information at a sentential level (as opposed to a simple form or the noun phrase level), and to the mode in which the assessment prompts were provided, or to a combination of these possibilities. The qualitative data underscore the importance of measuring the construct of unawareness from different sources—that is, both online and offline—given that processes demonstrated during different phases of exposure to the input may not be fully reported at the nonconcurrent stage of retrieval of such awareness. The hybrid research design that included both quantitative and qualitative analyses subsequently provided rich sources of data to establish the presence or absence of this construct (including levels of awareness), unexpected strategies used by learners while performing the assessment tasks, and deeper insights into learners' processes.

Empirically addressing the construct of awareness (or lack thereof) and its role in learning presents methodological challenges for the operationalization, the type of measurement, and the stage at which this construct is measured. The most researchers can attempt to do is carefully gather as much data as possible from multiple sources and stages, including both online and offline measures and different assessment tasks, in an effort to shed more light on the role of this construct. Further investigation of the role the construct of awareness plays in learning is clearly warranted.

(*Received 15 September 2009*)

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

NOTES

1. Hybrid research designs employ both qualitative and quantitative analyses of elicited data, promote higher internal validity of the study, and reduce the potential for committing a Type I or Type II error. A Type I error falsely reports a difference in the data when in fact there is none, whereas a Type II error reports that there is not a difference in the data when in fact there is one (Isaac & Michael, 1997).

2. The relatively high rate of exclusion of participants is typically reported in many studies (e.g., Leow, 1997, 1998a, 1998b; Rosa & Leow, 2004a, 2004b; Rosa & O'Neill, 1999) that have employed concurrent data elicitation procedures (e.g., verbal reports) to ascertain what processes are employed by learners while exposed to the L2 data. Concurrent data tend to reveal that many participants do not comply with the instructions they receive or do not perform according to the condition designed by the experimenter. It may be argued that eliminating these participants from the study for noncompliance raises the level of internal validity in the study.

3. Thanks to our statistician, who ran our analyses and provided the chance correction formula to compare items obtained on two- versus four-option tasks.

4. Paired-samples *t* tests conducted on the trained versus new items comparisons for the CACD and IACD production responses did not reveal any significant result, $t(33) = 1.291$, $p = .206$ and $t(33) = -1.391$, $p = .174$, respectively.

REFERENCES

Bialystok, E. (1979). Explicit and implicit judgments of L2 grammaticality. *Language Learning*, *29*, 81–103.

Bowles, M. A. (2003). The effects of textual input enhancement on language learning: An online/offline study of fourth-semester Spanish students. In P. Kempchinsky & P. Pineros (Eds.), *Theory, practice, and acquisition: Papers from the 6th Hispanic Linguistic Symposium and the 5th Conference on the Acquisition of Spanish and Portuguese* (pp. 359–411). Somerville, MA: Cascadilla Press.

Bowles, M. A. (2004). L2 glossing: To CALL or not CALL. *Hispania*, *87*, 541–552.

Bowles, M. A. (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. A., & Leow, R. P. (2005). Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation. *Studies in Second Language Acquisition*, *27*, 415–440.

Curran, T., & Keele, S. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 189–202.

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*, *113*, 541–555.

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1985). On consciousness in syntactic learning and judgment: A reply to Reber, Allen, and Regan. *Journal of Experimental Psychology: General*, *114*, 25–32.

Ellis, R. (2001). Investigating form-focused instruction. *Language Learning*, *51*, 1–46.

Ellis, R. (2004). The definition and measurement of explicit knowledge. *Language Learning*, *54*, 227–275.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.

Isaac, S., & Michael, W. B. (1997). *Handbook in research and evaluation: A collection of principles, methods, and strategies useful in the planning, design, and evaluation of studies in education and behavioral sciences* (3rd ed.). San Diego, CA: Educational and Industrial Testing Services.

Leow, R. P. (1995). Modality and intake in SLA. *Studies in Second Language Acquisition*, *17*, 79–89.

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning*, *47*, 467–505.
Leow, R. P. (1998a). The effects of amount and type of exposure on adult learners' L2 development in SLA. *Modern Language Journal*, *82*, 49–68.
Leow, R. P. (1998b). Toward operationalizing the process of attention in second language acquisition: Evidence for Tomlin and Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics*, *19*, 133–159.
Leow, R. P. (2000). A study of the role of awareness in foreign language behavior: Aware versus unaware learners. *Studies in Second Language Acquisition*, *22*, 557–584.
Leow, R. P. (2001). Attention, awareness and foreign language behavior. *Language Learning*, *51*, 113–155.
Leow, R. P. (2006). The role of awareness in L2 development: Theory, research, and pedagogy. *Indonesian Journal of English Language Teaching*, *2*, 125–139.
Leow, R. P., & Bowles, M. A. (2005). Attention and awareness in SLA. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 179–206). Washington, DC: Georgetown University Press.
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.
Litman, L., & Reber, A. S. (2005). Implicit cognition and thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 431–453). New York: Cambridge University Press.
Nissen, M., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1–32.
Perruchet, P., & Amorim, M.-A. (1992). Conscious knowledge and changes in performance in sequence learning: Evidence against dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 785–800.
Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or fragmentary knowledge? *Journal of Experimental Psychology: General*, *119*, 264–275.
Perruchet, P., & Pacteau, C. (1991). Implicit acquisition of abstract knowledge about artificial grammar: Some methodological and conceptual issues. *Journal of Experimental Psychology: General*, *120*, 112–116.
Perruchet, P., & Pacton, S. (1996). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*, 233–238.
Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional sets. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 88–94.
Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford: Oxford University Press.
Robinson, P. (1995). Aptitude, awareness, and the fundamental similarity of implicit and explicit second language learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 303–357). Honolulu: University of Hawai'i Press.
Rosa, E., & Leow, R. P. (2004a). Awareness, different learning conditions, and L2 development. *Applied Psycholinguistics*, *25*, 269–292.
Rosa, E., & Leow, R. P. (2004b). Computerized task-based exposure, explicitness, type of feedback, and Spanish L2 development. *Modern Language Journal*, *88*, 192–216.
Rosa, E., & O'Neill, M. D. (1999). Explicitness, intake, and the issue of awareness. *Studies in Second Language Acquisition*, *21*, 511–556.
Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning*, *57*, 165–199.
Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, *29*, 67–100.
Sachs, R., & Suh, B.-R. (2007). Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In A. Mackey (Ed.), *Conversational interaction in second-language acquisition: A series of empirical studies* (pp. 197–227). Oxford: Oxford University Press.
Sanz, C., Lin, H.-J., Lado, B., Bowden, H. W., & Stafford, C. A. (2009). Concurrent verbalizations, pedagogical conditions, and reactivity: Two CALL studies. *Language Learning*, *59*, 33–71.
Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129–158.

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206–226.
Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, *11*, 11–26.
Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu: University of Hawai'i Press.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
Shanks, D. R., Green, R. E. A., & Kolodny, J. A. (1994). A critical examination of the evidence for unconscious (implicit) learning. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 837–860). Cambridge, MA: MIT Press.
Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*, 367–447.
Shook, D. J. (1994). FL/L2 reading, grammatical information, and the input-to-intake phenomenon. *Applied Language Learning*, *5*, 57–93.
Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, *16*, 183–203.
VanPatten, B., & Sanz, C. (1995). From input to output: Processing instruction and communicative tasks. In F. Eckman, D. Highland, P. Lee, J. Mileham, & R. Rutkowski Weber (Eds.), *Second language acquisition: Theory and pedagogy* (pp. 169–185). Mahwah, NJ: Erlbaum.
Williams, J. N. (2004). Implicit learning of form-meaning connections. In J. Williams, B. VanPatten, S. Rott, & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition* (pp. 203–218). Mahwah, NJ: Erlbaum.
Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, *27*, 269–304.
Wong, W. (2003). Textual enhancement and simplified input: Effects on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning*, *13*, 109–132.

APPENDIX A

ADDITIONAL SENTENCES USED IN THE LEARNING TASK

1. I pushed forward gi turtle which was struggling to go over a rock.
2. I almost stepped on gi turtles while walking on the beach.
3. I was riding gi horse when I saw our guest.
4. I took pictures of gi horses lining up in front of me.
5. I stopped the boy from throwing a rock to ul rabbit across the street.
6. From a distance, it is difficult to distinguish ul rabbits from squirrels.
7. The children wanted to touch ul elephant across the field.
8. I could hear ul elephants marching in the far field.
9. I turned off ro lamp on my night stand before I went to sleep.
10. The sales person said ro lamps in front of us were all on sale.
11. I dropped ro spoon I was holding in my hand.
12. I polished ro spoons for our dinner party.
13. The girl went to the dining room to blow out ne candle on her birthday cake.
14. From my apartment window, I could see ne candles lit in a nearby park.

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press

15. I heard ne phone ringing in the next room.
16. The sales person struggled to reach ne phones on the top of the display shelf.

APPENDIX B

SENTENCES USED IN THE MC AND PRODUCTION ASSESSMENT TASKS

MC Task (some sentences were modified to clarify the distance)

1. While lying in the field the woman was bothered by gi bees.
2. While at the nursery the baby slipped off ro stool that she was standing on.
3. The wolf wanted to eat ul birds on the tree.
4. Before the battle the soldier checked the time on ro clock he was holding.
5. The waitress was told to go and clear ne table at the end of the room.
6. The zookeeper struggled to control gi monkey on his shoulder.
7. The thieves searched the living room for ne vases stored in the attic.
8. The hunter spotted ul lion with binoculars.

Production Task

1. At the animal farm, the child held gi rabbit.
2. I saw ne spoons on the kitchen floor from the living room.
3. From the beach, I saw ul turtle swimming in the ocean.
4. I was surprised when ro phone I was about to use rang.
5. I couldn't see ul horses on the field without glasses.
6. While I was walking by gi elephants, I was surprised to see how big they were.
7. I asked my sister to turn off ne lamp in the next room.
8. I placed ro candles on my friend's birthday cake.

https://doi.org/10.1017/S0272263110000045 Published online by Cambridge University Press