Applied Psycholinguistics **25** (2004), 269–292
Printed in the United States of America
DOI: 10.1017.S0142716404001134

Awareness, different learning conditions, and second language development

ELENA M. ROSA and RONALD P. LEOW
Georgetown University

ADDRESS FOR CORRESPONDENCE
Elena M. Rosa, 4642 Kearney Court, Alexandria, VA 22311.

ABSTRACT
This study investigated the role of awareness in different learning conditions in relation to second language (L2) development. Adult learners of Spanish were exposed to past conditional sentences in one of five versions of a computerized problem-solving task that vary in their degree of explicitness. By employing both on-line and off-line data elicitation procedures, two main issues were addressed: whether exposure to L2 input under different conditions had a differential impact on learners' awareness and whether different levels of awareness influenced learners' ability to recognize and produce new exemplars of the target structure immediately after exposure to the input and over time. Results indicated that higher levels of awareness were not only associated with more explicit conditions but were also substantially more effective than lower levels.

It is generally accepted in cognitive psychology, cognitive science, and in the second language acquisition (SLA) field that attention is a necessary condition for learning (e.g., Robinson, 1995a; Schmidt, 2001; Shiffrin & Schneider, 1977; Tomlin & Villa, 1994). Differing from the general consensus regarding the role of attention in second/foreign language (L2) learning, there is much less agreement among researchers with respect to the role played by awareness. On the one hand, some experimental studies in cognitive psychology have suggested a dissociation between awareness and learning (e.g., Curran & Keele, 1993; Nissen & Bullemer, 1987; Reber, 1976, 1993). On the other hand, several scholars in cognitive psychology and in SLA have argued that learning without awareness, if at all possible, is so limited that its effects are negligible (e.g., Dulany, Carlson, & Dewey, 1984; Robinson, 1995a, 1995b; Schmidt, 1995).

The role of awareness in relation to attention in L2 development has been discussed from a theoretical perspective. Schmidt's (e.g., 1990, 1993, 1995) noticing hypothesis proposes that noticing, that is, attention accompanied by some low level of awareness, is the "necessary and sufficient condition for converting input to intake" (1990, p. 129). In other words, according to Schmidt, the only material that can be taken in is that of which the individual is aware. To account for item learning versus system learning, Schmidt posits two levels of awareness: awareness at the levels of noticing and understanding. This view has been contradicted by

© 2004 Cambridge University Press 0142-7164/04 $12.00

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

Tomlin and Villa (1994), whose fine-grained model of attention makes detection (that is, attention without any crucial role for awareness) the first step toward language development. In their view, awareness may enhance input processing, but its presence is not required. Robinson (1995b) incorporates both Schmidt's and Tomlin and Villa's attentional postulations into his model by positing that noticing includes detection and rehearsal in short-term memory. By placing noticing further along the acquisitional process than detection, Robinson concurs with Schmidt that lack of awareness precludes learning.

EMPIRICAL STUDIES ON THE ROLE OF AWARENESS IN SLA

To address the roles of attention and awareness in L2 development, it is important that SLA investigations include and operationalize these constructs. Recently, several studies have employed either postexposure questionnaires (e.g., Robinson, 1996) or think-aloud protocols (e.g., Leow, 1997, 1998a, 1998b, 2000, 2001; Rosa & O'Neill, 1999). The latter studies methodologically established that reported attention was paid to targeted morphological and syntactic forms/structures in the L2 data before analyses were conducted to address the effects of attention and awareness. For example, Leow (1997, 2001) requested participants to think aloud while performing an experimental task. Subsequent coding of these nonmetalinguistic protocols revealed three levels of awareness: noticing, reporting, and rule formation.

Studies employing concurrent data elicitation procedures found positive support for the role of awareness in L2 development. More specifically, these studies reported that (a) awareness at the levels of noticing and understanding contributed substantially to a significant increase in learners' ability to take in (Leow, 1997, 2001; Rosa & O'Neill, 1999) and produce in writing (Leow, 1997, 2001) the targeted form or structure, (b) awareness at the level of understanding led to significant more intake (Leow, 1997, 2001; Rosa & O'Neill, 1999) and written production (Leow, 1997, 2001) when compared to awareness at the level of noticing, (c) there appears to be a correlation between awareness at the level of understanding and usage of hypothesis testing and rule formation (Leow, 1997, 2001; Rosa & O'Neill, 1999), and (d) there appears to be a correlation between level of awareness and formal instruction or directions to search for a rule (Rosa & O'Neill, 1999). In other words, more awareness appeared to translate into more intake and learning.

Studies employing on-line data elicitation procedures are crucial in that they are the only SLA investigations that have put forward a psycholinguistically motivated fine-grained account of different levels of awareness based on concurrent or on-line assessment of learners' cognitive processes (cf. Leow, 1999, 2000, for further discussion). However, these studies have not addressed delayed effects of awareness nor whether learners are capable of systematizing the grammatical information they became aware of while processing the L2 data.

LEARNING CONDITIONS AND L2 DEVELOPMENT

Several studies in SLA have examined the effects of different types of learning conditions on L2 development (e.g., DeKeyser, 1995, 1997; Ellis, 1993; Robinson,

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

1995a, 1996, 1997a, 1997b; Rosa & O'Neill, 1999). Following the field of cognitive psychology, the distinction between implicit and explicit learning in SLA studies has typically been operationalized by creating (a) a so-called *explicit learning condition* in which learners are instructed to look for rules underlying the input and (b) a so-called *implicit learning condition* in which individuals are instructed to memorize the input. The premise underlying these studies is that explicit learning occurs when individuals are instructed to search for rules whereas implicit learning appears when learners are told to memorize.

Within the learning condition research strand in SLA, only one study, to our knowledge, has incorporated specific features in the experimental task in an effort to investigate the role of awareness in L2 development (Rosa & O'Neill, 1999). The 67 learners who participated in this study were randomly divided into five conditions of different degrees of explicitness. These conditions were created by varying two factors, namely, explicit instruction in Spanish contrary to fact conditional sentences, on the one hand, and directions to search for rules on the other. In order to investigate the effects of awareness on intake, the researchers asked learners to think aloud while performing the problem-solving tasks. They found that (a) the degree of explicitness had a differential effect on intake; (b) the higher the level of awareness, the stronger the effect on intake; and (c) the conditions under which the task was administered influenced the way information was processed. However, whether awareness played a role in system learning or had a delayed effect on learners' performance was not addressed in this study.

The general findings of SLA studies suggest that explicit learning conditions may be positive for SLA in that the levels of accuracy developed under them seem to be higher than those developed under implicit learning conditions. At the same time, there is need to carefully gather data that do reflect the type of learning process assumed to take place in the experimental conditions created by the researcher (cf. Leow, 2000, for further discussion of the so-called representative nature of participants' performance in experimental cells in SLA research). In addition, this strand also needs to assess the product of learning (i.e., knowledge) by means of assessment methods other than grammaticality judgment tests (for example, tests measuring intake and production of a target structure).

Of special interest for the present study is the role of different types of feedback in learners' linguistic accuracy in specific learning conditions. In cognitive psychology, feedback has traditionally been linked to the process of hypothesis formation and testing (Estes, 1989), a process found in SLA studies investigating different levels of awareness (e.g., Leow, 1997, 2001; Rosa & O'Neill, 1999). One recent study (Sanz & Morgan–Short, 2004) investigated the effects of different types of computer-delivered feedback (implicit vs. explicit) on the acquisition of Spanish word order by first- and second-year speakers of English with some prior knowledge of Spanish word order. Although the authors reported that feedback did not play a crucial role when compared to exposure to structured input, participants' level of awareness was not methodologically addressed. The issue of different types of feedback in a learning condition context in order to find out which conditions of explicitness during input processing are more beneficial for the development of formal awareness and learning clearly warrants exploration.

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

The present study investigated the relationship between different conditions of exposure to L2 input and the levels of awareness raised during performance of problem-solving tasks under those conditions. It also investigated whether the levels of awareness reported during performance of tasks differing in their degree of explicitness had a differential effect on learners' ability to generalize their intake and production of a Spanish structure to new exemplars of that structure. The construct degree of explicitness was operationalized by manipulating the amount of grammatical information that learners received on the target structure, either in the form of feedback simultaneous to task performance or in the form of a grammatical explanation prior to the task. The variable awareness was operationalized and measured by means of think-aloud protocols collected concurrently to input processing. It was also measured off-line by means of a questionnaire administered immediately after completion of the experimental tasks.

The research questions guiding this study were the following:

1. Is there a relationship between different conditions of exposure to L2 input through tasks of varying degrees and types of explicitness and the level of awareness raised during performance on those tasks?
2. Does the level of awareness reported during performance of tasks of varying degrees and types of explicitness have a differential effect on learners' ability to generalize their intake of a Spanish structure to new exemplars of that structure as measured by a multiple-choice recognition test? If so, will the effects be maintained over a period of 3 weeks?
3. Does the level of awareness reported during performance of tasks of varying degrees and types of explicitness have a differential effect on learners' ability to generalize their knowledge to new exemplars of a Spanish structure as measured by a controlled written production test? If so, will the effects be maintained over a period of 3 weeks?

METHOD

Participants

Participants were 100 out of an original total of 132 adult native speakers of English enrolled in fifth semester Spanish courses (i.e., Spanish Advanced I) at an American university. To maximize the homogeneity of the sample, participants met the following criteria: (a) they were native speakers of English, (b) they spoke only English and did not receive more than 2 years of formal study in any other Romance language, (c) they demonstrated very little or no ability to recognize and no ability to produce the targeted structure, and (d) they demonstrated the same understanding of the temporal framework that corresponds to the different types of conditional sentences in English. Most of the 32 participants eliminated from the study failed to complete all parts of the study.

The targeted structure

The targeted structures were Spanish contrary to fact conditional sentences in the past. Conditional sentences can be divided into two main types, factual and contrary to fact. Factual conditional sentences imply experience whereas contrary

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

to fact conditional sentences imply lack of experience with the events expressed by both main and subordinate clauses (Campos, 1993). Contrary to fact conditional sentences may refer to future, present, or past time frames. Conditional sentences are considered to be problematic, highly complex structures in the context of both first (L1) and L2 acquisition (e.g., López Ornat, 1994), and are typically not taught in the first 2 years of language study.

For this study, two different types of contrary to fact conditional sentences in Spanish were chosen, namely, those that refer to the present or future with a result also in the present (e.g., *"Si Joe estudiara alemán, entendería el discurso de Kohl."* "If Joe studied German, he would understand Kohl's speech.") and those that refer to the past with a result in the present (e.g., *"Si Joe hubiera estudiado alemán, ahora entendería el discurso de Kohl."* "If Joe had studied German, he would now understand Kohl's speech."). The sequence of verb tenses for conditional sentences that refer to the present or future is imperfect subjunctive in the subordinate clause and conditional tense in the main clause. The sequence of verb tenses for conditional sentences that refer to the past is past perfect subjunctive in the subordinate clause and conditional tense in the main clause. The former type of conditional sentence (and within it, the use of the imperfect subjunctive) was used as a base structure that would be explicitly taught to all participants in order to form a baseline knowledge of Spanish conditional sentences that was common to both experimental and control groups. The latter type of structure, that is, contrary to fact conditional sentences in the past (and within it, the use of the past perfect subjunctive) was the real focus of this investigation.

Task features manipulated in the study

Selection of the task parameters in this study as independent variables was primarily based on their potential to help participants become aware of the formal structure of the L2 input presented within the tasks themselves. These parameters were task essentialness (±essential; Loschky & Bley–Vroman, 1993), provision of immediate feedback intrinsic to the task (±feedback, further manipulated with + explicit or + implicit; Loschky & Bley–Vroman, 1993), and presence of a pretask consisting of explicit information on the target grammatical structure (±explicit pretask; e.g., Foster & Skehan, 1996). The parameters were combined in the following ways to create five experimental task-based learning conditions and one control learning condition. An overview of the task features manipulated in the present study is displayed in Table 1.

The experimental tasks

To maximize the comparability between all conditions created for this study, exposure to the L2 input took place in a computer-based setting by means of a series of problem-solving tasks involving manipulation of the input. Participants in the control condition were also exposed to input in a computer setting but they did not perform a problem-solving task. All treatment materials for the experimental and control conditions were created using LIBRA (version 1.2) software for Macintosh (Farris, 1995).

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

Table 1. *Overview of task features manipulated in the present study*

	Features		
Tasks	Essentialness	Explicit Pretask	Feedback
EPEFE	+	+	+ (explicit)
EPIFE	+	+	+ (implicit)
EFE	+	−	+ (explicit)
EP	+	+	−
IFE	+	−	+ (implicit)
Control	−	−	−

Note: EPEFE, explicit pretask + explicit feedback; EPIFE, explicit pretask + implicit feedback; EFE, explicit feedback; EP, explicit pretask; IFE, implicit feedback; CO, control.

Table 1 shows that all experimental tasks were characterized by the feature of +essentialness, which contrasted with the control task marked with the minus (−) value of every one of the three target features in order to create as implicit a learning condition as possible. The design also allowed for the comparison of the effects of explicit grammatical information (+pretask) and feedback (explicit vs. implicit). Explicit grammatical information on the use and structure of Spanish conditional sentences provided in the EPEFE, EPIFE, and EP tasks was provided by several LIBRA grammar cards, all accompanied by glossed samples. Practice via an interactive multiple-choice question format was provided on four cards in which participants were asked to read a conditional sentence and figure out its time frame by clicking on one of two options: either present/future or past. Participants were given 8 min (based on the findings of a pilot test) to complete the pretask phase of the experiment.

Participants also received two types of feedback: explicit (EPEFE, EFE) and implicit (EPIFE, IFE) located at the bottom of the card where there was a button that read "click here for FEEDBACK" to be clicked before going on to the next page. For the explicit conditions, if the answer was correct, participants received a prompt reinforcing the reason why the participant's choice was the right one plus an explanation of how the targeted structure works in Spanish. If the answer was wrong, participants received a prompt saying that their choice was incorrect together with an explanation of the reason why and information on how the targeted structure works. In this case, participants were instructed to try again, so that they made the correct selection of time frame for each particular sentence before going on to the next card. For the implicit condition, participants only received feedback that indicated whether the answer was right or wrong.

Participants in the most explicit learning condition (EPEFE) received explicit grammatical explanation of the rule before performing the task and explicit feedback while performing the task. Participants in the EPIFE group also received the same grammatical information before the task but received implicit feedback during the task. The EFE group did not receive prior grammatical information but did receive explicit feedback during the task. On the other hand, the EP

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

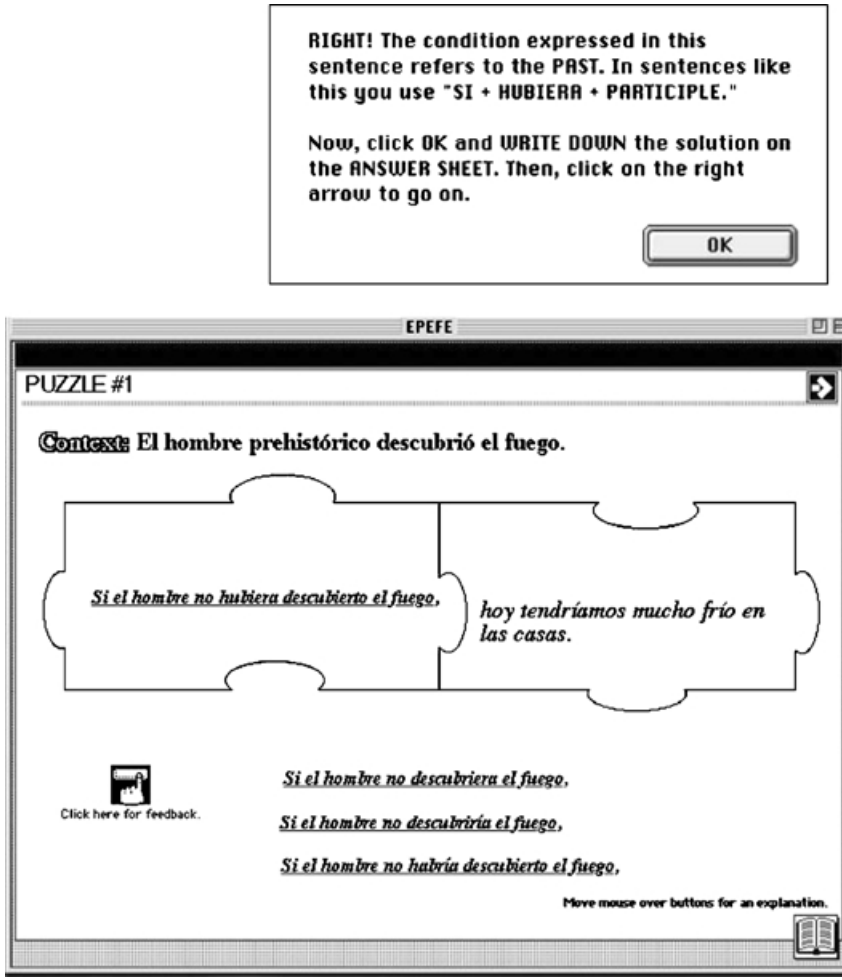


Figure 1. A sample puzzle.

participants did receive the explicit grammatical information before the task but no feedback during the task. Only implicit feedback was provided to the IFE participants while participants in the control condition read the same experimental sentences for meaning without any prior grammatical information or feedback.

The explicit pretask was immediately followed by the main problem-solving task. This part consisted of a series of 28 interactive LIBRA cards, each containing a Spanish contrary to fact conditional sentence presented in the form of a jigsaw puzzle. There were 10 contrary to fact conditionals in the present (i.e., *imperfect subjunctive + simple conditional*) and 18 contrary to fact conditional sentences in the past (i.e., *past perfect subjunctive + simple conditional*). In order to clarify the time frame of the sentences, each one was introduced by a context (see Figure 1

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

for samples). The 28 treatment sentences were randomly ordered for presentation in the form of problem-solving puzzle tasks in the case of the experimental groups and in the form of a sentence-reading activity in the case of the control group.

The participants' task was to solve all 28 puzzles, each of which had the following structure: there were two puzzle pieces on the computer screen, a main clause piece, and, right next to it, an empty piece. Underneath the two puzzle pieces there were four moveable subordinate clauses; with the only difference between them being the tense of the verbs. Participants were informed in bold that only one of the fragments was appropriate to complete the puzzle. This feature contributed to the essentialness of the task in that participants were primed to focus their attention on the target structure (both on the morphology of the four candidate verbal forms and on the function of the correct form, that is, on the reason why some tenses worked in certain contexts while not in others). Participants solved each puzzle by filling the empty piece with the different movable subordinate clauses and finding out which of these four clauses corresponded to the given main clause piece (see Figure 1 for sample puzzles).

In an effort to control for potential vocabulary problems related to the L2 input presented to participants during their respective treatments, two measures were adopted. First, care was taken to include only high-frequency vocabulary on the treatment sentences. Second, a button called *Dictionary* appeared on the bottom right-hand corner of the computer screen. By clicking on that button, the participants were able to look up any word that they did not understand. As initially predicted based on the pilot tests conducted prior to this study, most participants did not need to use the Dictionary feature at all. Out of the 100 participants in this experiment, only 7 used it, and in every case only one word was looked up.

In order to establish that attention was paid to the target structure, every time participants completed one sentence an alert dialog box appeared on the computer screen asking them to write down the sentence fragment that solved every puzzle on an answer sheet. In addition, the presence and levels of awareness raised during completion of the experimental task were assessed by means of think-aloud protocols (adapted from Rosa & O'Neill, 1999) collected concurrently to task performance. Based on a pilot test, participants were allowed 28 min to complete the tasks. Participants also completed a brief questionnaire aimed at further assessing their level of awareness of the target structure (Appendix A). Drawing from Robinson's investigations (e.g., 1996, 1997a), this questionnaire asked participants whether they were looking for patterns while completing the task and whether they could state any rules they might have come up with that could explain how the puzzle pieces worked. However, differing from Robinson's methodology, the questionnaire was administered immediately after the treatment, rather than after the assessment tests. Think-aloud protocols and postexposure questionnaires were administered to participants in all of the conditions, experimental and control, included in this study.

Assessment tests

To ensure unfamiliarity with the targeted structure, all participants completed two paper and pen pretests comprising a multiple-choice recognition test and a

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

controlled written production test. To prevent the multiple-recognition test items from providing additional input that could affect participants' performance on the production posttest, this task was administered after the production task.

Recognition tests

Three multiple-choice recognition tests (a pretest, an immediate posttest, and a delayed posttest) were used in the present study. The pretest consisted of 24 sentences, of which 10 were contrary to fact conditionals in the past (i.e., the targeted structure) and 14 were contrary to fact conditionals in the present/future, which in this investigation were treated as distracters.

Participants received instructions to decide which of four options best completed each sentence. The missing element was always the verb of the subordinate clause, which could be placed either in initial or in final position of the conditional sentence. The time allowed for this task was 8 min, and participants were encouraged to work as quickly as possible (see Appendix B for a sample of the pretest). Finally, the test started and ended with two distracters.

The posttest's structure was very similar to that of the pretest. However, the number of tokens of the targeted structure was increased from 10 to 20. Participants were given 11 min (based on the findings of a pilot test) to complete the recognition tests. The delayed posttest was the same as the first posttest but its items were randomly redistributed.

Controlled-production tests

In order to test whether participants had processed the L2 input presented in the treatment beyond the initial stage of intake, three written controlled-production tests (a pretest, an immediate posttest, and a delayed posttest) were administered.

The pretest consisted of 22 sentences, distributed as 10 target items and 12 distracters. For every item of this test, participants were given the main clause and the beginning of the subordinate clause of a conditional sentence. A verb appeared in parenthesis followed by a blank space. Participants were asked to complete the sentence logically by providing an appropriate verb form. This test had to be completed in a maximum time of 8 min (see Appendix C for a sample of the pretest), based on the results of a pilot test.

As was the case with the recognition tests, items were randomly ordered, the first and last two items were distracters, and there were two versions of this test with different ordering of sentences.

The immediate and delayed posttests were the same as the pretest, except that the amount of targeted items was increased from 10 to 20 to encompass a further distinction between old and new exemplars of the target structure. Participants were given 11 min to complete this test, based on the results of a pilot test.

Procedure

To ensure that participants were homogeneous in relation to the targeted structure, prior to the experiment, all participants received the same instruction on

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

Spanish contrary to fact conditional sentences in the present (i.e., *si + imperfect subjunctive + conditional*). A series of written activities were performed at home and corrected in class the following day. Immediately after, participants completed the background questionnaire as well as the recognition and written production pretests.

The data collection phase took place in a language laboratory equipped with Macintosh computers and tape recorders connected to microphones. Each session was preceded by a familiarization pre-training session that included a short practice session on the participants' assigned computers. After the familiarization session, participants were randomly assigned to one of the following conditions: *Group EPEFE*, *Group EPIFE*, *Group EFE*, *Group EP*, *Group IFE*, and *Group CO*.

Participants were asked to think aloud while completing their respective treatments. Immediately after, they completed a packet containing the awareness questionnaire, the controlled written production test, and the multiple-choice recognition test.

Finally, 3 weeks after the treatment, participants completed a delayed controlled written production test, a multiple-choice recognition test, and a debriefing questionnaire aimed at ensuring that participants had the same understanding of the time frames of English contrary to fact conditional sentences in the present, past and future. Figure 2 summarizes the experimental design.

Scoring procedures

Multiple-choice recognition tests. In this test, 1 point was given to each correct target sentence and 0 to each incorrect one. Every participant obtained a global score on the pretest (out of a maximum score of 10 points), and two scores on the immediate and delayed posttests, one for old target items (i.e., sentences previously encountered during the treatment phase of the experiment) and one for novel exemplars of the target rule. The maximum possible posttest score was 10 for old items and 10 for new ones.

The threshold for admission to the study was 3 correct answers on the pretest out of a maximum of 10 to cover for those cases in which correct answers may be selected by chance.

Controlled-production tests. In the sentence-completion tests, 1 point was given to each target item in which the participant used past perfect subjunctive and 0 when a different tense was used. Errors of agreement and of verb form were not taken into account as long as it was evident that the participant intended to use a past perfect subjunctive.

Every participant received a global score on the pretest (out of a maximum of 10 points) and two scores in the immediate and on the delayed posttest respectively, one for old target items and one for novel ones. The maximum possible score on the posttests was 10 for both types of target items.

In order to be admitted to the study, participants' scores on the controlled production pretest had to be 0 out of 10 points. This way, any effects found could be safely attributed to the treatments rather than to previous knowledge that participants may have had prior to the study.

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

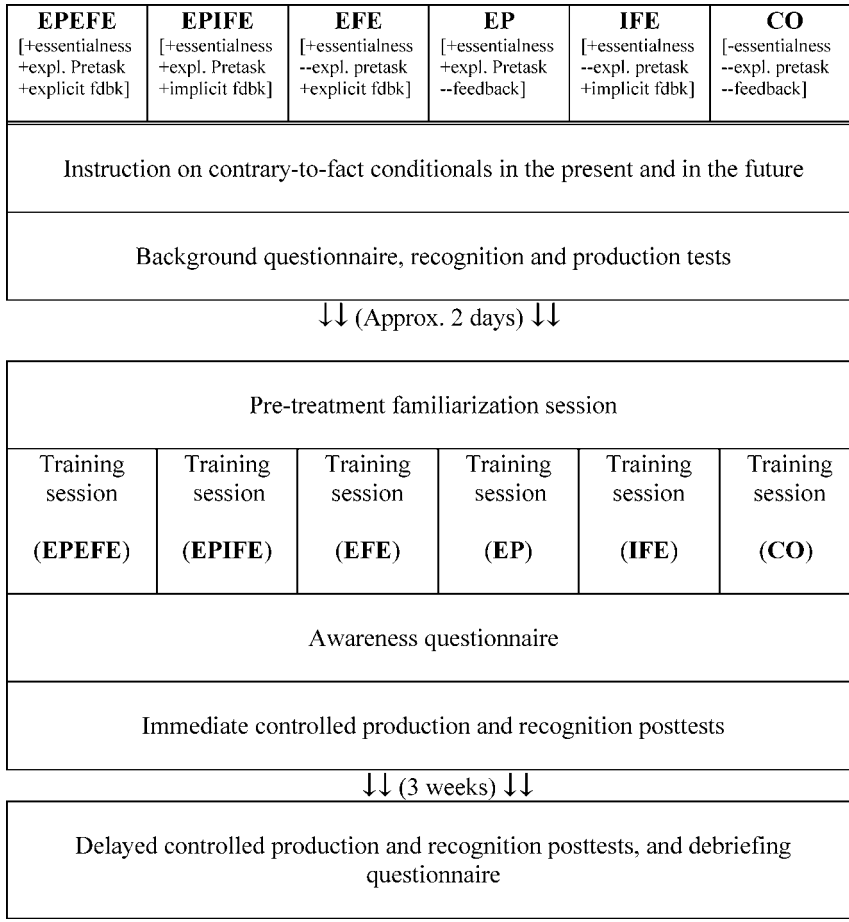


Figure 2. An experimental design.

Measures of awareness: Think-aloud protocols and postexposure questionnaires

Two types of measures of awareness were included in this study, namely, think-aloud protocols produced concurrently to input processing and a written questionnaire administered immediately after exposure. The former was considered the main assessment instrument for awareness (after all, its on-line nature made it more reliable as a source of cognitive data), and all participants, regardless of their randomly assigned treatment group, were reclassified according to the type of think-aloud protocol they produced.

The think-aloud protocols reported by participants in this investigation were coded closely following Rosa and O'Neill's (1999) method. In other words, in line with Schmidt's (e.g., 1990) levels of awareness, verbal reports were classified

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

as [Noticing] ([N]) or [Understanding] ([U]). In addition, those cases in which participants failed to provide evidence of being aware of the target form, both in their verbal protocols and in their postexposure questionnaires, were coded as [No Report] or [NR]. Postexposure questionnaires were used as a method of ensuring that (a) individuals in the [N] category did not possess a rule that could account for the use of the targeted structure and (b) individuals in the [NR] category had a second chance to demonstrate whether they had developed any level of awareness of the targeted structure.

Two native speakers of Spanish (including one of the researchers) independently read and coded a sample of 20 protocols randomly chosen from the 100 obtained in this investigation. Raters received a basic definition of each level of awareness (similar to the definitions presented in previous studies, e.g., Rosa & O'Neill, 1999, but without any specific examples of protocols). On the basis of those definitions, they classified each report as either [NR], [N], or [U]. Agreement was 100% and the remaining 80 verbal reports were coded by the researcher without the help of the second rater.

RESULTS

To answer research question 1, which addressed the relationship between the independent variables Learning Condition (Control, EFE, EP, EPEFE, EPIFE, IFE) and Awareness Report, chi squares were performed on frequency counts of each type of awareness report [NR], [N], and [U] per learning condition, using Microsoft Excel 5.0 with an alpha level set at .05. A considerable number of cells, however, presented expected frequencies lower than 5, which made it necessary to collapse both variables into fewer levels.

Given that collapsing variables inevitably causes loss of information, the various levels of Learning Condition and Awareness were collapsed in four different ways, each of which provided additional data that contributed to a richer global picture of the relationship between the two variables. Hence, four chi squares were conducted.

The first chi-square analysis was the most general one, because it collapsed the factor Learning Condition into only two levels, [Control] and [Experimental] (which included conditions EFE, EP, EPEFE, EPIFE, and IFE). Awareness Report was also reduced to two levels, namely, [Report] (which included categories [N] and [U]) and [NR]. The first chi-square analysis for Learning Condition by Awareness Level revealed a significant relationship ($df = 1, \chi^2 = 48.415, p < .001$).

For the second chi square, Learning Condition was collapsed into three levels: feedback only (i.e., EFE, IFE), EP + feedback (i.e., EPEFE, EPIFE), and EP only. The Control condition with only two reports of noticing was left out of the chi-square analysis. With respect to the factor Awareness, its original three levels were reduced to two, low (i.e., [N]) and high (i.e., [U]). The category [NR] was not included, given that only 2 out of the 83 participants in experimental conditions presented this type of reports. The second chi-square analysis for Learning Condition by Awareness Level also revealed a significant relationship ($df = 2, \chi^2 = 10.25, p < .01$).

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

Table 2. *Recognition of new items: Means table for awareness report by time*

	Count	Mean	Std. Dev.	Std. Error
Pretest				
[U]	59	2.220	0.911	0.119
[N]	28	2.179	1.278	0.242
[NR]	13	1.892	1.316	0.365
Posttest				
[U]	59	8.288	2.060	0.268
[N]	28	5.250	2.952	0.558
[NR]	13	2.000	2.646	0.734
Delayed posttest				
[U]	59	7.610	2.282	0.297
[N]	28	4.821	2.957	0.559
[NR]	13	3.154	2.460	0.960

For the third chi square, the factor Learning Condition was collapsed yet another way, whereas Awareness Report was treated in the same form as for the second chi square. In order to capture the concept of explicitness, Learning Condition was collapsed according to the number of sources of explicit information on the target structure available to participants during task completion: two sources (i.e., feedback + EP), one source (i.e., either feedback or EP), and no sources. The third chi-square analysis for Learning Condition by Awareness Level also revealed a significant relationship ($df = 2$, $\chi^2 = 11.264$, $p < .01$).

For the final chi-square analysis, the factor Awareness Report was reduced to the same two categories detailed above, that is, high ([U]) and low ([N]). Learning condition was manipulated two different ways. On one hand, the category Control was eliminated from the chi-square analysis in order to avoid cells with less than five observations. On the other hand, the levels EPEFE and EPIFE were collapsed into one category, once again, for the purpose of avoiding low count cells. Like the results above, the fourth chi-square analysis for Learning Condition by Awareness Level revealed a significant relationship ($df = 3$, $\chi^2 = 11.865$, $p < .01$).

The significant results found for all four chi-square analyses suggest that the two variables at hand (i.e., Learning Condition and Awareness Report) were significantly related.

As previously explained, in order to address the effects of the levels of awareness reports identified for this study, subjects were redistributed into three groups, namely [U], [N], and [NR]. The mean scores and standard deviations for Awareness Report by Time for the recognition and written production of new exemplars are reported in Tables 2 and 3, respectively. A one-way analysis of variance (ANOVA) conducted on raw pretest recognition scores revealed no significant differences between the groups ($F = 1.889$, $p = .1568$). With respect to production, between-group comparability on the pretest was obvious because all participants had a score of 0. Therefore, it could be safely assumed that score changes between test sessions were due to the effect of the independent variable (e.g., Awareness Report) rather than to previous knowledge of the targeted structure.

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

Table 3. *Production of new items: Means table for awareness report by time*

	Count	Mean	Std. Dev.	Std. Error
Pretest				
[U]	59	0.000	0.000	0.000
[N]	28	0.000	0.000	0.000
[NR]	13	0.000	0.000	0.000
Posttest				
[U]	59	6.949	2.649	0.345
[N]	28	3.179	3.044	0.575
[NR]	13	0.769	2.774	0.769
Delayed posttest				
[U]	59	4.559	3.435	0.447
[N]	28	2.536	2.715	0.513
[NR]	13	0.538	1.198	0.332

Table 4. *Recognition of new items: ANOVA for Awareness Report (AR) by Time (T)*

Source	*df*	Sums of Squares	Mean Squares	*F*	*p*
AR	2	549.959	274.980	37.725	.0001*
T	2	461.434	230.717	68.527	.0001*
T × AR	4	238.194	59.548	17.687	.0001*
Residual	194	653.160	3.367		

*$p < .05$.

In order to address research questions 2 and 3, separate 3 × 3 repeated-measures ANOVAs were conducted on participants' scores in recognition and production of new exemplars of the target structure. Each ANOVA used a one between-subject, one within-subject design. The between-subject factor was Awareness Report and the within-subject factor was Time. Each ANOVA was also followed by a Scheffé post hoc test and means comparisons. Finally, individual ANOVAs and means comparisons were also performed for each group in order to assess its respective score growth or loss across test sessions. The ANOVAs were conducted on the SuperANOVA statistical software program with the alpha level set at .05.

Research question 2 addressed the effects of Awareness Report on learners' ability to recognize instances of the targeted structure not encountered during the treatment. The analysis showed a significant main effect for Awareness Report, a significant main effect for Time, and a significant interaction between Awareness Report and Time (see Table 4). The main effect found for Awareness Report must be interpreted in light of the interaction of this factor with Time, which is shown in Figure 3.

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

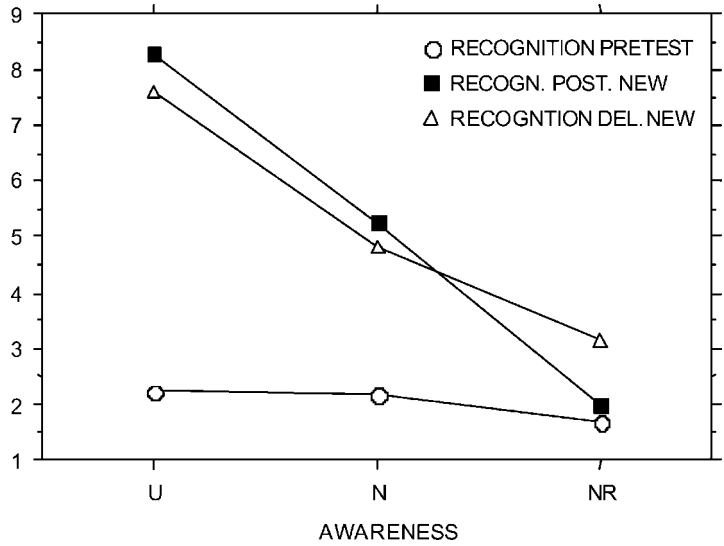


Figure 3. Awareness Report mean scores on recognition of new items across Time.

Table 5. *Recognition of new items: Post hoc Scheffé for Awareness Reports*

				d (Effect Size)	
	Diff.	Crit. Diff.	*p*	Posttest	Delayed
NR vs.					
N	1.801	1.301	.0037*	1.137	0.594
U	3.757	1.187	.0001*	2.911	1.927
N vs.					
U	1.956	0.889	.0001*	1.297	1.117

*$p < .05$.

As Figure 3 clearly shows, the interaction between Awareness Report and Time was due to the fact that the groups varied to a different extent across test sessions. In other words, the between-group variation at the time of the posttest was not the same as the between-group variation at the time of the delayed posttest.

A post hoc Scheffé test (Table 5), conducted to identify which contrasts were responsible for the main effect found for Awareness Report, revealed that [NR] varied significantly from [N] and [U], and [N] varied significantly from [U]. On both the immediate and delayed posttests, large effect sizes were observed between [U] and [N] ($d = 1.297$ and 1.117, respectively) and between [U] and [NR] ($d = 2.911$ and 1.927, respectively). A large effect size was also observed between [N] and [NR] on the immediate posttest ($d = 1.137$), and a medium effect size was found on the delayed posttest between these two groups ($d = 0.594$).[1]

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

Table 6. *Production of new items: ANOVA for Awareness Report (AR) by Time (T)*

Source	*df*	Sums of Squares	Mean Squares	*F*	*p*
AR	2	474.470	237.235	27.579	.0001*
T	2	482.725	241.363	57.412	.0001*
T × AR	4	280.630	70.158	16.688	.0001*
Residual	194	815.590	4.204		

*$p < .05$.

Separate repeated-measures ANOVAs with a zero between-subject, one within-subject design were then conducted on each group's raw scores for recognition of new items. The results revealed that Time had a significant main effect in the case of groups [N] and [U] ($p = .0001$ for both), but not in that of [NR] ($p = .1878$). Post hoc means comparisons showed significant score increases from the pretest to the posttest in the case of both [N] ($F = 132.071$, $p = .0001$) and [U] ($F = 435.132$, $p = .0001$). Similarly, significant differences from the pretest to the delayed posttest were present for [N] ($F = 20.280$, $p = .0001$) and [U] ($F = 343.328$, $p = .0001$). Finally, a significant score loss between the posttest and the delayed posttest was found for [U] ($F = 5.432$, $p = .0215$), but not for [N] ($F = 0.533$, $p = .4684$).

The results summarized above suggest a positive answer to research question 2, that is, the level of awareness reported through the combination of the on-line verbal protocols and postexposure questionnaires appeared to have a differential effect on learners' ability to recognize new items, and these differential effects tended to be maintained after a period of 3 weeks.

Research question 3 addressed the immediate and delayed effects of the factor Awareness Report on learners' production of exemplars of the target structure not encountered during the treatment phase. Results showed significant main effects for Awareness Report and for Time and a significant interaction between the two factors. These results are displayed in Table 6. The main effect found for Awareness Report must be interpreted in light of the interaction of this factor with Time, which is clearly shown in Figure 4.

A post hoc Scheffé test was performed to identify which group contrasts would account for the main effect of Awareness Report. Results were similar to those obtained for research question 2: [NR] varied significantly from [N] and [U], and [N] varied significantly from [U] (Table 7). On both the immediate and delayed posttests, large effect sizes were observed between [U] and [NR] ($d = 2.315$ and 1.318, respectively) and between [N] and [NR] ($d = 0.814$ and 0.889, respectively). A large effect size was also observed between [U] and [N] on the immediate posttest ($d = 1.359$) and a medium effect size was found on the delayed posttest between these two groups ($d = 0.631$).

Separate repeated-measures ANOVAs with a zero between-subject, one within-subject design were then conducted on each group's raw scores for production of new items. Similarly to the results obtained for recognition of new items, the effect

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

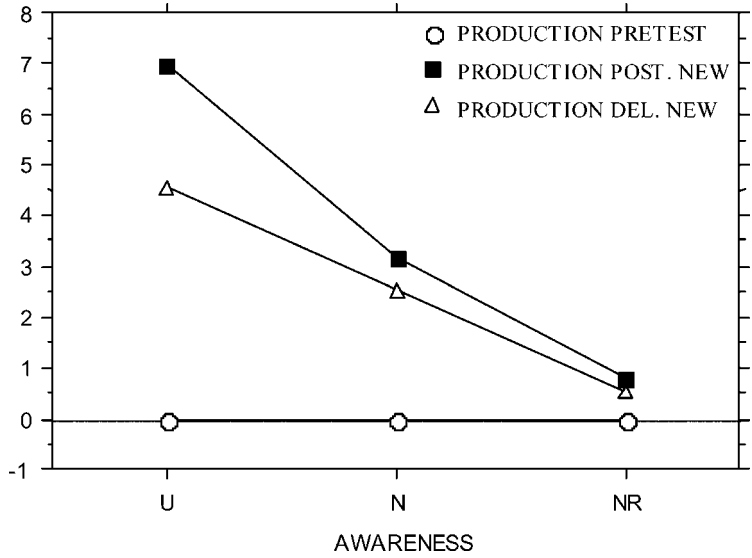


Figure 4. Awareness Report mean scores on production of new items across Time.

Table 7. *Production of new items: Post hoc Scheffé for Awareness Reports*

				d (Effect Size)	
	Diff.	Crit. Diff.	*p*	Posttest	Delayed
NR vs.					
N	1.469	1.413	.0396*	0.814	0.889
U	3.400	1.290	.0001*	2.315	1.318
N vs.					
U	1.931	0.966	.0001*	1.359	0.631

*$p < .05$.

of the factor Time was only significant for groups [N] and [U] ($p = .0001$). Post hoc means comparisons revealed that both [N] and [U] improved significantly from the pretest to the immediate posttest ($F = 36.988$, $p = .0001$, and $F = 295.544$, $p = .0001$, respectively). For the pretest to the delayed posttest, the same pattern was found, that is, for [N] ($F = 23.540$, $p = .0001$), and for [U] ($F = 127.221$, $p = .0001$). Finally, scores experienced a significant decrease from the immediate to the delayed posttest in the case of [U] ($F = 34.954$, $p = .0001$), but not in that of [N] ($F = 1.513$, $p = .2240$).

These analyses suggest that the various levels of awareness identified for this investigation had a differential effect on learners' production of new items, both immediately after exposure to the treatment and 3 weeks after that.

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

DISCUSSION

The significant results of the four chi squares suggest a positive answer to research question 1, that is, there was a positive relationship between the explicitness of the various learning conditions (EPEFE, EPIFE, EFE, EP, IFE, Control), and the levels of awareness reported by learners in each condition ([U], [N], [NR]). The results suggest that, in general, learners processed input with various degrees of awareness of the target structure as a function of the explicitness of each given learning condition. First, learners exposed to L2 input in a non-task-based way failed to demonstrate awareness at the level of understanding, whereas all other conditions had at least some learners who were able to verbalize rules accounting for the target structure. Second, for those learners exposed to the L2 input in a task-based way, more sources of information on the target structure translated into larger frequencies of reports of awareness at the level of understanding. When only one source of information was available to learners, techniques such as a pretask or explicit feedback were more effective than implicit feedback in helping learners develop higher and more sophisticated levels of awareness. Third, only two learners in the experimental conditions failed to report awareness of the target structure, be it at the level of noticing or at that of understanding, which points toward the efficacy of the task chosen for the present investigation as a way of raising learners' awareness and facilitating the extraction of linguistic information out of the L2 input included in them.

The results concur to a large extent with those of Rosa and O'Neill (1999), who found that learners who had received explicit information on their target structure had a clear tendency to report higher levels of awareness than individuals who were exposed to the L2 input in less explicit ways. The inclusion of postexposure questionnaires as a second method of establishing levels of awareness contributed to the strengthening of the reliability of the conclusions regarding learners' cognitive processes.

Research questions 2 and 3 investigated the issue of item learning versus system learning, that is, whether different levels of awareness raised during L2 input processing have a differential impact on learners' ability to generalize their intake (research question 2) or written production (research question 3) of a Spanish structure to new exemplars of that structure and whether these effects are maintained over a period of 3 weeks.

As the results reported in Tables 4 and 5 show, the various levels of awareness identified for the present investigation appeared to have had a differential impact on learners' ability to generalize their intake to new exemplars of the target structure. Learners who demonstrated awareness at the level of understanding had significantly higher scores than learners who demonstrated noticing or who gave no evidence of awareness in their verbal and written reports. At the same time, learners reporting awareness at the level of noticing experienced a score growth significantly larger than learners in the [NR] group, findings also found in Leow (1997, 2001) and Rosa and O'Neill (1999). These results differ from Robinson (1995a, 1996, 1997b), who found that only awareness at the level of understanding had a significant impact on learning; in this study noticing also had positive effects on learners' performance. This discrepancy can be justified in light of the

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

different nature of the assessment tests used in each one of these investigations: whereas Leow, Rosa and O'Neill, and the current investigation incorporated a task-based approach and tested learners' performance in a series of recognition tests measuring intake (i.e., cognitive registration), Robinson's learners were exposed to input in a non-task-based way and participants' performance was assessed through grammaticality judgment tests aimed at measuring learning.

The findings also coincide to a large extent with those of Leow (1997, 2001), who, also in the context of a problem-solving task but at the level of item learning, investigated whether and how different levels of awareness influenced learners' recognition and controlled production of a series of morphological forms in Spanish. In Leow's study, the behavior of learners who demonstrated higher levels of awareness (or what the author calls meta-awareness) had two main characteristics. Qualitatively, they used conceptually driven processing such as hypothesis testing and morphological rule formation; quantitatively, they performed significantly better both on recognition and on controlled production posttests.

With respect to the issue of time, the results displayed in Figure 3 reinforce the hypothesis of the beneficial effects of high levels of awareness on learners' ability to generalize their intake to new items. Both the [N] and the [U] categories demonstrated a significant score improvement from the pretest to the immediate posttest. In addition, no significant score decrease was observed for the [N] group between the immediate and the delayed posttest. The [U] group, however, did experience a significant score loss over time, but this possible negative effect of time was counterbalanced by the fact that [U] learners still managed to maintain a significant score gain between the pretest and the delayed posttest. Finally, the [NR] category behaved as predicted, that is, no significant differences were observed between the mean scores obtained in the pretest, immediate, and delayed posttest. On the basis of these results, it is possible to conclude that the differential effects of awareness on learners' ability to recognize new exemplars of the target structure were generally maintained over a period of 3 weeks.

In line with the findings of Leow (1997, 2000, 2001) and Rosa and O'Neill (1999), the present investigation supports the idea that higher levels of awareness are associated with more complex and sophisticated types of input processing that allow for extraction of patterns from specific instances of a given L2 structure. The process of hypothesis formation and testing manifested itself with different intensity depending on the learning conditions under which the computerized problem-solving tasks were performed. Those conditions providing learners with implicit or explicit feedback simultaneous to task completion were more effective in activating this processing strategy (EPEFE, EPIFE, EFE IFE). In contrast, only 5 learners reported hypothesis formation and testing out of the total of 16 included in the EP group (the only experimental condition providing no feedback to individuals on the accuracy of their responses while processing input). Finally, none of the learners in the control condition formulated hypotheses on how the target structure works in Spanish.

The findings reported above provide empirical support to the claim that feedback (simultaneous or not to input processing) prompts learners to analyze the L2 input in an effort to extract generalizations that can be used in other contexts (i.e., Carroll, Swain, & Roberge, 1992; Tomasello & Herron, 1988). Traditionally, the

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

association between feedback and hypothesis formation and testing was largely based on inferential data provided by postexposure assessment tests. In that respect the present investigation represents an important contribution, since it provides empirical evidence of the types of cognitive processes activated through feedback during on-line input processing.

The results (Tables 5 and 6) also suggest a positive answer to research question 3 that addressed learners' ability to produce exemplars of the target structure that had not been encountered during the treatment phase of the experiment. Similar to the findings for recognition of new items, awareness at the level of understanding manifested itself in significantly higher production scores than awareness at the level of noticing. Moreover, learners who reported having noticed the target structure performed significantly better than those who did not report being aware of it either in the think-aloud protocols or in the subsequent questionnaires.

Similar results were also reported by Leow (1997, 2001), who observed a downward trend in accuracy of production scores as a function of decreasing levels of awareness (from 87% accuracy observed in learners reporting meta-awareness of the target morphological forms to 66% in learners not reporting such meta-awareness). Delayed posttest results showed the same trend, with mean scores ranging from 4.56 for [U] learners to 2.54 for the [N] category and 0.54 for [NR]. Once again, these results provide evidence of the facilitative role of awareness in L2 development as espoused by Schmidt (e.g., 1990), Leow (1997, 2000, 2001), and Robinson (1995a, 1996, 1997b).

With reference to the delayed effects of awareness, the results suggest similar conclusions to those reached for recognition of new items. First, the factor of Time had no main effect in the case of the [NR] group, which showed no significant improvement from the pretest to the immediate or to the delayed posttest. Given that 11 of the 13 learners included in this category belonged to the Control condition, it can be claimed that exposure to L2 input in a non-task-based way demanding comprehension of content was not an effective method of helping learners attain high levels of awareness and obtain generalizable knowledge that could be used to produce novel exemplars of the target structure.

Second, significant main effects were found for the factor Time in the data for [N] and [U]. The production patterns found for these two groups were identical to their respective recognition patterns, that is, they both experienced a significant score increase from the pretest to the immediate posttest. Furthermore, whereas learners in the [N] category showed no substantial score loss from the immediate to the delayed posttest, those in the [U] condition performed significantly worse over time. Nonetheless, the apparent detrimental effects of time on [U] learners' production ability, were counterbalanced by the fact that the differences between the pretest and the delayed posttest were still significant. On the basis of these results it can be concluded that, in the context of the task and the target structure chosen for the present investigation, the higher the level of awareness attained during input processing the stronger its effect on learners' production accuracy, both immediately after exposure and overtime. In addition, time had very limited detrimental effects, because learners in the [N] and [U] groups maintained a significant score gain 3 weeks after the treatments had been administered.

All in all, in the context of the computerized problem-solving tasks created for this experiment, awareness appeared to have a differential effect on learners'

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

controlled written production of a target Spanish structure. Generally, higher levels of awareness were reflected in superior accuracy of production and in fewer cases of overgeneralization. Finally, the positive effects of awareness tended to be maintained for 3 weeks after exposure to the L2 input had taken place.

Limitations and further research

The first limitation concerns the assessment tests used in the current study. Even though, in an effort to tap both intake and production, two different types of posttests were used (recognition and controlled written production) their restricted nature makes the conclusions that can be drawn from them rather restricted. As a possible expansion of the present investigation, it would be interesting to see whether the observed patterns of performance for learners in each of the conditions would be the same in more complex assessment tasks (e.g., free written production).

The second limitation is rooted in the short time allowed for the treatment phase of this experiment. Time constraints prevented the researchers from administering the treatment in various sessions, which, at least in theory, could have affected L2 development beyond recognition and controlled written production. In this sense, future studies could reinforce the treatment through various sessions spread over a longer period of time. The consequences of such a prolonged treatment in terms of free oral and written production could then be investigated.

The third limitation relates to the short time that elapsed between the immediate and delayed posttests. Future investigations ought to examine the issue of awareness through their long-term effects. Once again, the length of the experimental treatments should be extended in order for their learning effects to be maintained over a long period of time.

Finally, the conclusions drawn above should be interpreted with caution and applied only to the specific context and learner population included in the present investigation. Additional research including different target structures, proficiency level learners, and even L2s is necessary before clear-cut generalizations can be drawn on the role of explicitness and awareness in L2 development.

CONCLUSION AND IMPLICATIONS

The results of this study have both theoretical and methodological implications. Theoretically, this study contributes to the discussion on the role of awareness in L2 development in relation to learning conditions that vary in their relative degrees and types of explicitness. More precisely, some of the postulates regarding the issue of awareness and how different levels may impact L2 development received additional empirical support in the present investigation. Methodologically, learners' cognitive processes during task completion were carefully assessed in order to determine whether the external conditions created by each given task prompted the types of processes that had been initially predicted by the researchers. This design follows that found in other recent studies that employ concurrent data elicitation procedures in an effort to address the internal validity of the study, especially with respect to the representative nature of learners' performances in experimental cells.

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

APPENDIX A
Postexposure awareness questionnaire for the experimental groups

Please read the following questions carefully regarding the task that you just completed. Answer them **IN ENGLISH** with as much detail as possible. You have as much time as you need. Please answer the questions **IN ORDER** and **DO NOT GO BACK**.

1. What do you think the purpose of this experiment is?

   _____

   _____

2. What do you think the purpose of the puzzle task was?

   _____

   _____

3. Do you think you know how the puzzle pieces worked, that is, which puzzle piece was the solution to each puzzle before using the feedback help?

   ___ **Yes, I know** ___ **No, I don't know** ___ **I think I know**

4. If you had to explain to someone how to solve the puzzles, could you give them some sort of a "rule" that would help them solve the puzzles without the help of the feedback button? If you could, please, write down that "rule."

   _____

   _____

   _____

APPENDIX B
Multiple-choice recognition pretest (version A)

The following activity contains 24 Spanish sentences of different types. You may or may not be familiar with some of these types of sentences. Read them and circle the letter that best completes each sentence. You will have to **work as quickly as possible** since you have only **8 minutes** to complete this task. You must work in order from 1 to 24 and **you may not go back to check your work.**

1. El autobús de Rosslyn a Georgetown tardaría menos de 10 minutos en llegar a la universidad si el tráfico en el Key Bridge no _____ tan terrible por las mañanas.
   a. fuera
   b. habría sido
   c. será
   d. hubiera sido
2. El programa de David Letterman tendría mucha más audiencia si _____ a las 9 de la noche en vez de a las 11.
   a. fuera
   b. fuera sido
   c. hubiera sido
   d. sería
3. Si el hombre prehistórico no _____ el fuego, hoy tendríamos mucho frío en las casas.
   a. hubiera descubierto
   b. descubriría
   c. habiera descubierto
   d. descubriera

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

APPENDIX C
Controlled-production pretest (version A)

The following activity contains 32 Spanish sentences of different types. You may or may not be familiar with some of these types of sentences. Read them carefully and complete the blank spaces logically by conjugating the verbs in parenthesis in the tense that you think is most appropriate in each case. You will have to **work as quickly as possible** since you have only **11 minutes** to complete this task. You must work in order from 1 to 32 and **you may not go back to check your work.**

1. Si las películas de la televisión no (SER) _____ por lo general tan violentas, los niños podrían verlas sin problema.
2. Solamente pasaría las vacaciones de Navidad en Washington DC si no (PODER) _____ ganar suficiente dinero en los próximos dos meses para comprar un billete de avión para ir a visitar a mi familia.
3. Si Einstein no (FORMULAR) _____ la Ley de la Relatividad, la ciencia no estaría tan avanzada hoy en día.
4. Ahora me podría ir de vacaciones a España si no (GASTAR) _____ el dinero que tenía guardado en el banco.
5. Luciano Pavarotti no sería tan famoso si no (CANTAR) _____ ópera tan maravillosamente bien.

NOTES

1. An effect size of .80 and above is considered a large effect. Effect sizes between .50 and .80 are viewed as medium effects, whereas values below .50 are considered small effects (e.g., Cohen, 1988; Norris & Ortega, 2000).

REFERENCES

Campos, H. (1993). *De la oración simple a la oración compuesta: Curso superior de gramática española.* Washington, DC: Georgetown University Press.

Carroll, S., Swain, M., & Roberge, Y. (1992). The role of feedback in adult SLA: Error correction and morphological generalizations. *Applied Psycholinguistics, 13*, 173–198.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

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

DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition, 17*, 379–410.

DeKeyser, R. (1997). Beyond explicit rule learning. *Studies in Second Language Acquisition, 19*, 195–221.

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

Ellis, N. (1993). Rules and instances in foreign language learning: Interactions of implicit and explicit knowledge. *European Journal of Cognitive Psychology, 5*, 289–319.

Estes, W. (1989). Learning theory. In A. Lesgold & R. Glaser (Eds.), *Foundations for a psychology of education* (pp. 1–49). Hillsdale, NJ: Erlbaum.

Farris, M. (1995). *Libra 2.1*. Southwest Texas State University, Division of Media Services.

Foster, P., & Skehan, P. (1996). The influence of planning time and task type on second language performance. *Studies in Second Language Acquisition, 18*, 299–323.

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

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press

Leow, R. (1998a). The effects of amount and type of exposure on adult learners' L2 development in SLA. *Modern Language Journal, 82*, 49–68.
Leow, R. (1998b). Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics, 19*, 133–159.
Leow, R. P. (1999). The role of attention in second/foreign language classroom research: Methodological issues. In F. Martínez–Gil & J. Gutiérrez–Rexach (Eds.), *Advances in Hispanic linguistics: Papers from the 2nd Hispanic Linguistics Symposium* (pp. 60–71). Somerville, MA: Cascadilla Press.
Leow, R. P. (2000). A study of the role of awareness in foreign language behavior: Aware vs. 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.
López Ornat, S. (1994). *La adquisición de la lengua española*. Madrid: Siglo XXI de España Editores.
Loschky, L., & Bley–Vroman, R. (1993). Grammar and task-based learning. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 123–167). Clevedon, UK: Multilingual Matters.
Nissen, M., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*, 1–32.
Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*, 417–528.
Reber, A. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 88–94.
Reber, A. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford: Clarendon Press.
Robinson, P. (1995a). 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 Hawaii Press.
Robinson, P. (1995b). Review article: Attention, memory and the noticing hypothesis. *Language Learning, 45* 283–331.
Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search and instructed conditions. *Studies in Second Language Acquisition, 18*, 27–68.
Robinson, P. (1997a). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition, 19*, 223–247.
Robinson, P. (1997b). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning, 47*, 45–99.
Rosa, E., & O'Neill, M. (1999). Explicitness, intake and the issue of awareness: Another piece to the puzzle. *Studies in Second Language Acquisition*, *21*, 511–566.
Sanz, C., & Morgan–Short, K. (2004). Positive evidence vs. explicit rule presentation and explicit negative feedback: A computer-assisted study. *Language Learning, 54*, 35–78.
Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–158.
Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206–226.
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 Hawaii Press.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127–190.
Tomasello, M., & Herron, C. (1988). Down the garden path: Inducing and correcting overgeneralization errors in the foreign language classroom. *Applied Psycholinguistics, 9*, 237–246.
Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*, 183–203.

https://doi.org/10.1017/S0142716404001134 Published online by Cambridge University Press