CrossMark

Annual Review of Applied Linguistics, **36** (2016), pp. 205–229.
© Cambridge University Press, 2016
doi: 10.1017/S0267190515000069

Task-Based Versus Task-Supported Language Instruction: An Experimental Study

SHAOFENG LI
University of Auckland
s.li@auckland.ac.nz

ROD ELLIS
University of Auckland and Shanghai International Studies University
r.ellis@auckland.ac.nz

YAN ZHU
Fudan University
yan_zhu@fudan.edu.cn

ABSTRACT

This study investigated the effectiveness of task-based and task-supported instruction in the acquisition of the English passive construction—a structure about which learners had limited prior knowledge. A total of 150 Chinese middle school English as a foreign language (EFL) learners were randomly assigned to five groups—one control group who only took the pretest and posttests and four experimental groups who attended a 2-hour treatment session where they performed two dictogloss tasks in groups, each including a reporting phase when the learners took turns to tell the narrative. Among the four experimental groups, one just performed the two oral tasks; a second group received explicit instruction before performing the tasks; a third group received within-task feedback but no explicit instruction; and the fourth group received both explicit instruction and within-task feedback. Treatment effects were gauged via a grammaticality judgment test (GJT) and an elicited imitation test (EIT). On the GJT, the conditions with explicit instruction and/or feedback led to significant gains with explicit instruction plus feedback showing the largest effects. On the EIT, there was no effect for any of the three treatment groups when the data were analyzed for the whole cohort. However, when the learners were subdivided into those with zero and some prior knowledge based on their pretest EIT scores, explicit instruction plus within-task feedback was more effective than the other treatment types for the latter.

INTRODUCTION

The utilization of tasks in language teaching or in the design of a research study involves decision making at a number of different levels:

1. Type of task (focused or unfocused)
2. Preparation for the performance of the task

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

3. Task performance
4. Posttask activity

The purpose of the study reported in this article is to investigate a number of the options involved at these different levels and the impact that they have on the learning of a specific grammatical structure.

Type of Task

An unfocused task is designed to elicit general samples of language use. In contrast, a focused task is designed to provide a communicative context for the use of a predetermined linguistic feature (or features). In general, pedagogic proposals have favored the use of unfocused tasks. Long (2015), for example, proposed that tasks be selected on the basis of a needs analysis to identify the "target tasks" relevant to a specific group of learners and suggested how "pedagogic tasks" can then be derived from the target tasks. Such tasks will generally be unfocused, as any single target task is unlikely to involve the use of a variety of linguistic forms. Skehan (1998) also argued for unfocused tasks on the grounds that task design should incorporate factors that influence the complexity, accuracy, and fluency of language use rather than the use of a predetermined linguistic feature.

There is, however, a pedagogic case for focused tasks. In many teaching contexts, teachers are tied to implementing a structural syllabus. Thus, if task-based language teaching (TBLT) is to work in such a context, teachers will need tasks that create contexts for the communicative use of specific linguistic (typically grammatical) features. Focused tasks can also be employed to help learners acquire fragile features, such as morphological inflections, that even successful learners often fail to acquire. Such tasks provide a context for intensive use of such features.

In research, too, there is a need for focused tasks in order to investigate whether the performance of tasks under different conditions results in learning. By preselecting a linguistic feature, experimental researchers can pretest learners' existing knowledge of that feature and then posttest to establish if any learning has resulted from the performance of the task. A number of studies (e.g., Doughty & Varela, 1998; Ellis, Loewen, & Erlam, 2006; Lyster, 2004; Mackey, 1999) have made use of focused tasks in this way.

There is, however, both a theoretical and technical problem with focused tasks. The theoretical problem is that of learner readiness. It can be argued that learners will only benefit from performing a focused task if they have reached a stage of development that permits them to acquire the targeted feature. While the inviolability of acquisition sequences is now a matter of some debate (see the Special Issue in *Language Learning*: Hulstijn, Ellis, & Eskildsen, 2015), there is also ample evidence to show that there are developmental constraints governing which grammatical features learners are able to learn (Pienemann, 2005). It is the difficulty in determining when learners are ready to acquire a particular grammatical feature that has led advocates of TBLT such as Long (2015) to reject the use of focused tasks.

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

The technical problem concerns the difficulty in designing tasks that obligate the production of a predetermined linguistic feature. Loschky and Bley-Vroman (1993) pointed out that whereas it is relatively easy to design structure-based production tasks that make it "natural" or "useful" to deploy a specific grammatical feature, it is much more difficult to make it "essential," and they concluded that "the goal in production tasks is likely to be limited to task-utility or task-naturalness" (p. 139). Much depends on the choice of target feature. Thus while task essentialness may be achievable for "questions" (as in Mackey, 1999), it is likely to be very difficult and perhaps impossible for those features whose meaning can be successfully conveyed without recourse to the grammatical forms that encode them in the target language. It is, for example, difficult to create task essentialness for the passive voice (see Williams & Evans, 1998), as learners can always opt to communicate using the active voice.

Focused tasks are used in both TBLT, where there is no prior explicit instruction, and in task-supported language teaching (TSLT), where there is. We consider these options in the following sections.

Preparing for the Performance of a Task

While learners can be asked to plunge immediately into the performance of task, considerable thought has been given to how learners can be prepared to perform a task so as to optimize its potential for acquisition. Skehan (1996) suggested two broad ways of achieving this—by assisting learners to deal with the cognitive demands of the task or by helping them with the language they needed to perform the task. Skehan's own research focused on the effect of strategic planning in helping learners to deal, in particular, with the cognitive demands of a task and what effect this has on the complexity, accuracy, and fluency of learner production when they perform the task. Here, however, we consider another option—the provision of linguistic guidance in the pretask stage of a lesson involving a focused task.

The most obvious way of providing such guidance for a grammar-focused task is direct explicit instruction. There is now ample evidence that explicit instruction is effective in enabling learners to use targeted features more accurately. A key finding of Norris and Ortega's (2000) and Spada and Tomita's (2010) meta-analyses of form-focused instruction studies was that explicit instruction is more effective than implicit instruction. However, not all the explicit treatments in the studies included in these meta-analyses were task-based, and many of the studies did not examine the effects of the instruction on spontaneous free production and thus did not show whether the explicit instruction resulted in changes in the learners' implicit/procedural knowledge. R. Ellis's (2007) narrative review claimed a benefit for explicit instruction even in studies that measured learning in free production tasks but Andringa, de Glopper, and Hacquebord (2011) concluded from their own review of studies that there was still little solid evidence to support the claim that explicit instruction is superior when learning is measured in a free production task.

Some advocates of TBLT reject providing learners with explicit instruction prior to performing a task. Long (2015), for example, while accepting that TBLT

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

"invokes a symbiotic combination of implicit and explicit learning" (p. 8), rejects any role for explicit instruction in TBLT while acknowledging that TSLT, which incorporates explicit instruction, may serve as a bridge between traditional synthetic syllabi and "genuine" task-based approaches. For Long, attention to form needs ideally to occur online while learners are performing a task in order to facilitate the form–function mapping central to acquisition. In other words, he argued for reactive rather than proactive focus on form (FonF). Others, however, have seen a role for proactive explicit instruction. Shehadeh (2005) championed TSLT for foreign language instructional contexts. Similarly, Sato (2010) argued that, in the Japanese context, tasks are best reserved for the free production stage in a presentation–practice–production (PPP) approach.

The key issue, then, is whether explicit instruction followed by the performance of a task results in better learning than simply performing the task by itself or with reactive FonF. From a theoretical perspective, the answer to this question depends on the model of memory systems one adopts. According to theories that are premised on separate neural systems for implicit and explicit knowledge (R. Ellis, 1994; Reber & Squire, 1998), the direct role of explicit knowledge in the acquisition of implicit knowledge is a minor one, largely limited to assisting learners to form an initial conscious representation of a target feature after which statistical learning processes automatically take over. In a single or integrated model of memory (Shanks, 2005), explicit knowledge transforms through practice into a state that allows for easy, automated performance. According to this model, explicit knowledge gained from explicit instruction is available for use through the course of learning and constitutes an integral part of the eventual expert knowledge system. Similarly, DeKeyser (1998), an advocate of skill-acquisition theory, argued that "proceduralization is achieved by engaging in the target behavior—or procedure—*while* temporarily leaning on declarative crutches" (p. 49, italics added).

Both models would seem to lend support to explicit instruction. In the case of a dual memory system, explicit instruction helps learners construct the initial conscious representation of the target structure, which then facilitates statistical learning processes that are implicit and rapidly take over. However, explicit instruction may be unnecessary as learners appear to automatically form their own initial explicit representation of a grammatical feature without assistance (see Morgan-Short et al., 2015). Furthermore, explicit instruction may even have a deleterious effect, as it can retard the development of skilled performance. In the case of a single, integrated mode, explicit instruction has a more important role to play, as the explicit representation it provides will be available for use by learners throughout the learning process. It provides the basis for subsequent proceduralization. In other words, the dual model suggests a task-based approach where explicit training may help but is not needed, as is claimed by the weak-interface position (N. Ellis, 2005; R. Ellis, 1994) while the single, integrated model (and also the skill-learning model) point to the advantage of a task-supported approach that includes a priori explicit instruction that caters for a transformation of explicit into implicit knowledge as in the strong-interface position.

To the best of our knowledge, there are no studies in second language acquisition (SLA) that have compared groups of learners who perform a focused task with or

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

without prior explicit instruction. There are, of course, plenty of studies that have compared implicit and explicit language instruction, but these studies typically involved other form-focused instruction options—such as exercises of a controlled nature (as in Andringa et al., 2011), integrated explicit instruction (as in Spada, Jessop, Suzuki, Tomita, & Valeo, 2014), corrective feedback while performing a task (Lyster, 2004; Sanz & Morgan-Short, 2004), or explicit instruction on completion of a task (Muronoi, 2000)—and/or they did not involve "tasks" as defined in the TBLT literature (e.g., Ellis & Shintani, 2014). The study that comes closest to the one we report below, which sought to compare groups that performed a task with or without a priori explicit instruction, is Williams and Evans (1998). They compared one group of university English as a second language (ESL) learners who were exposed to an input flood providing frequent exposure to the target structures (participial adjectives of emotive verbs and passive constructions) with another group that received the same input preceded by explicit instruction. There was also a control group that completed just the pre- and posttests. Both groups outperformed the control group on both a sentence-completion and narrative task for the use of passive verb forms, but there were no differences between the two experimental groups. This study, then, suggests that a priori explicit instruction confers no advantage. However, the learners in both of the experimental groups in this study already possessed explicit knowledge of passives at the start and so were not clearly distinguished in terms of their explicit knowledge. In the study we report below, we investigate learners with either very little or no prior knowledge of the target structure.

In the domain of psychology, there are studies that have investigated whether pretraining instruction has any added effect over purely implicit learning and so addressed "whether explicit knowledge information contributes directly to learning to perform or whether it reflects concomitant knowledge in a separate representational system that is epiphenomenal to task performance" (Sanchez & Reber, 2013, p. 3). Using an implicit perceptual-motor sequence learning task, Sanchez and Reber (2013) compared a group that received explicit pretraining instruction with a group that simply performed the perceptual learning task. In this study, both groups had no existing explicit knowledge of the sequence at the beginning of the study. The key finding was that although the participants developed robust explicit knowledge of the sequence as a result of the explicit training, this did not lead to a better performance or better learning than those participants who did not receive any training. They concluded that the learning that took place in both groups was implicit and purely the result of practice. However, Sanchez and Reber acknowledged that explicit knowledge can play a role in initially guiding performance, a view reflective of the weak-interface positon.

In an SLA study, Cintrón-Valentín and Ellis (2015) investigated the effect of providing pretraining (consisting of both explicit instruction in one group and verb pretraining where the learners translated first language [L1] verb forms into the second language [L2] in another group) and text enhancement on *ab initio* learners of Latin in a third group. The aim of this study was to investigate whether the pretraining enabled them to attend to morphological cues of temporality instead

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

of relying on adverbial markers. Cintrón-Valentín and Ellis reported that all three treatments resulted in increased attention to the morphological cues as the learners processed stimuli containing both adverb and morphological markers of temporality, reversing the natural bias to focus on the more salient and reliable adverb cues. All three treatments led to acquisition of verbal morphology as measured by both comprehension and production posttests and were broadly equivalent in their effect. This laboratory-based study demonstrates the effectiveness of focusing learners' attention on form as they process input. However, the practice phases involved discrete sentence stimuli, so the learners were not asked to process input while their primary attention was focused on achieving a task outcome, as in task-based or task-supported teaching.

There is clearly a need to investigate whether prior explicit instruction has an add-on effect to the learning that results from performing a task. As noted above, there are divergent opinions, with some arguing for "pure" TBLT involving online FonF and others favoring TSLT that includes a pretask stage involving explicit instruction. These differing pedagogic positions correlate with different theoretical models, with TBLT more clearly compatible with a dual model that posits separate neural implicit and explicit systems and TSLT drawing on a model that views memory as a unitary, integrated system or on skill-learning theory. However, to date, there has been no study involving tasks that has adequately addressed these competing positions.

Task Performance

Irrespective of whether there is any prior explicit instruction, opportunities arise to attract learners' attention to form—both implicitly and explicitly—while a task is being performed. Indeed, a central tenet of TBLT is that, while purely implicit learning is possible, drawing attention to form–meaning mappings will facilitate learning and, in the case of "fragile" grammatical features or features where the L1 has a blocking effect, may be necessary. It was this tenet that underlay the importance that Long (1991) attached to FonF in task-based teaching. Initially, FonF was closely linked to the interaction hypothesis (Long, 1996), according to which the communication problems that arise when learners are performing tasks trigger *the negotiation of meaning*, providing learners with input and pushing them to modify their output and thereby fostering form–function mapping. Subsequently, researchers such as Lyster (e.g., Lyster, 2001), pointed out that FonF need not be limited to occasions when communication problems occur but can be deliberately engineered whenever a linguistic problem arises in what has become known as *negotiation of form*. Deliberately engineering attention to form sits easily with focused tasks that have been designed to elicit the use of a predetermined linguistic feature.

Focus on form can occur preemptively when either the teacher explicitly mentions a linguistic feature surmised to be problematic to learners or when learners themselves take time out from performing a task to enquire about the use of a linguistic feature (Ellis, Basturkmen, & Loewen, 2001), but in general,

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

reactive FonF consisting of corrective feedback has been seen as of greater importance. There is now a wealth of research that has investigated corrective feedback in a task-based context. Meta-analyses (Li, 2010; Lyster & Saito, 2010) provide conclusive evidence that correcting learners' errors while they perform tasks has a strong effect on subsequent accurate use of grammatical features in both controlled and free production. Arguments in favor of specific corrective strategies have abounded, leading, in particular, to debates about the efficacy of recasts (see Goo & Mackey, 2013; Lyster & Ranta, 2013). Recasts are typically viewed as an implicit corrective feedback strategy, but they can also be quite explicit, especially if combined with some other strategy as in "corrective recasts" (Doughty & Varela, 1998). In these, the learner's erroneous utterance is first repeated, and then, if the learner fails to self-correct, a recast is provided. Given that the research points to an advantage of more explicit types of feedback (Ellis et al., 2006), corrective recasts would seem to have the greater potential to facilitate acquisition. Corrective recasts were the strategy used in the study reported below.

A number of corrective feedback studies have investigated the joint effect of pretask explicit instruction and one or more types of feedback strategies. Lyster (2004) provided form-focused instruction prior to the learners performing a series of communicative tasks during which they also received feedback. Pedagogically, this makes good sense as arguably it enhances the likelihood of the treatment having an effect on acquisition. Lyster reported clear effects for both recasts and output-prompting feedback on learners' acquisition of French gender marking. However, such a design does not distinguish the separate (and potentially different) effects of proactive and reactive FonF that we wish to investigate.

Many other corrective feedback studies (e.g., Yang & Lyster, 2010) have simply investigated the effects of feedback without any preliminary explicit instruction. These show that when tasks are performed and corrective feedback is provided, gains in accurate use of the target structures occur. It should be noted, however, that in these studies, the learners were in possession of both explicit and some procedural/implicit knowledge of the target structures prior to the treatment. In other words, these studies, like the vast bulk of classroom-based studies that have investigated corrective feedback, only provide evidence that task-based instruction that includes corrective feedback enables learners to *consolidate* their existing grammatical knowledge. The question remains whether such instruction is effective for structures that learners have no or very little existing prior knowledge. One study that has addressed this (Li, Zhu, & Ellis, in press) reported that feedback in the context of performing tasks resulted in learning but only of explicit knowledge as measured by a grammaticality judgment test (GJT). This study failed to find any significant effect for the instruction on the learners' procedural/implicit knowledge as measured by an elicited imitation test (EIT). The study reported below builds on Li et al. (in press) by examining the effect of corrective feedback in conditions where one group of learners received only feedback when performing the tasks and another group received pretask explicit instruction along with corrective feedback.

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

Posttask Activity

A further possibility is that any attention to form is delayed until a task has been completed. This is the position favored by some teachers and teacher educators on the grounds that focusing on form while a task is being performed—either preemptively or reactively—will "interfere" with the primary purpose of a task, namely, to develop communicative confidence and fluency. For this reason, some teacher guides (e.g., Hedge, 2000) propose that teachers observe their students performing tasks, make notes about any linguistic problems they manifest, and then address these explicitly when the task is over. Basturkmen, Loewen, and Ellis (2004), however, found that such beliefs (or advice) were not acted on by the teachers they observed, who typically went ahead with focusing on form during the performance of tasks although often without being aware of doing so.

SLA theories point to the advantage of immediate FonF on the grounds that it encourages learners to notice linguistic features in a context where these features are of communicative importance to them and to carry out the cognitive comparisons that promote change in interlanguage systems (Doughty, 2001). Transfer-appropriate processing also supports immediate rather than delayed FonF. Drawing on this theory, which states that "we can use what we have learned if the cognitive processes that are active during learning are similar to those that are active during retrieval" (p. 27), Lightbown (2008) suggested that learners are more likely to transfer what they have learned if attention to form is integrated into an ongoing communicative activity rather than if it occurs in isolation as is the case when it is delayed. However, there is clear a need to investigate empirically the relative effects of immediate and delayed FonF in task-based teaching, but to date, this remains largely unstudied. In the study reported below we did not include a posttask FonF condition, but see Li et al. (in press).

Measuring Acquisition

R. Ellis (2006) suggested that acquisition can mean three things: (a) internalization of a new linguistic form, (b) increased control over a partially acquired feature, and (c) progress along an acquisition sequence. In the study reported below, we are concerned with (a), which we examined in terms of whether the instructional treatments led to the *emergence* of the target structure in learners who demonstrated no knowledge of it prior to the instruction, and (b) which we examined in terms of whether there were *gains in accuracy* of the target structure as a result of the instruction in learners who demonstrated some prior knowledge of the structure.

In measuring both emergence and gains in accuracy, we investigated the learners' use of the target structure by means of a GJT and an EIT, with a view to establishing whether the instruction led to explicit/declarative and/or to implicit/procedural knowledge. We drew on research by R. Ellis (2005), which provided evidence that these two types of tests provide relatively separate measures of these two types of knowledge. Further studies (e.g., Bowles, 2011; Zhang, 2015) have supported this claim. However, doubts have been raised as to whether the EIT constitutes a valid test of implicit knowledge. Suzuki and DeKeyser (2015), for example, reported a study that concluded that the EIT is better viewed as measuring

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

automatized explicit knowledge. However, what does not seem to be controversial is that this test affords a measure of automated knowledge—whether explicit or implicit. As DeKeyser (2003) pointed out, automated explicit knowledge can be viewed as functionally equivalent to implicit knowledge. Thus, we frame our study as addressing whether four instructional conditions—Task Only, Explicit Instruction + Task, Task + Corrective Feedback, and Explicit Instruction + Task + Corrective Feedback—differed in the effect they had on the learners' ability to judge ungrammatical sentences (serving as a measure of their explicit/declarative knowledge) and on their ability to imitate sentences containing both grammatical and ungrammatical exemplars of the target structure (serving as a measure of automated knowledge).

RESEARCH QUESTIONS

The study was designed to investigate what effect different conditions pertaining to the performance of two focused tasks had on learners' explicit/declarative and automated knowledge of the target structure (past passive). The four conditions were:

- Task only: Performance of the task without any intervention. We argue that this corresponds to Long and Robinson's (1998) focus on meaning (FonM) condition.
- Performance of the tasks following pretask explicit instruction of the target structure, as in TSLT.
- Performance of the tasks with corrective feedback directed at the target structure (i.e., Long & Robinson's FonF and pure TBLT).
- Performance of the tasks with both pretask explicit instruction and corrective feedback (i.e., a stronger version TSLT).

We were also interested in whether there was any difference in the effect of these different conditions depending on whether the learners had no initial knowledge of past passive or had already acquired some (limited) knowledge of it. To this end, we formulated two research questions:

1. Are there any differences in the effect of the different conditions on learners' acquisition of (a) explicit/declarative knowledge and (b) automated knowledge of past passive?
2. Do the effects of the different conditions differ depending on whether the learners had no or some prior knowledge of past passive?

METHOD

Participants

The participants were 150 eighth-grade EFL learners at a middle school in southeastern China. The learners were 13–15 years old, with the average age of 14.10 years ($SD = .54$), and with an average of 6.18 years previous study of English. Five intact classes were randomly selected from a total of 18 classes at the eighth grade, each comprising 55 to 60 students. Out of each of the five classes, 30 students were

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

randomly selected to form five participant groups. The five groups were equivalent in their general proficiency represented by their midterm exam grades, $F(4, 143) = .18$, $p = .95$; and in their pretest scores on the GJT, $F(4, 140) = .23$, $p = .92$; as well as on the EIT, $F(4, 138) = .32$, $p = .86$. On a weekly basis, the learners attended seven 45-minute English lessons which were heavily grammar-based and emphasized the importance of the mastery of linguistic knowledge for test preparation purposes. Overall, these learners were homogeneous in terms of age, L1 background, L2 proficiency, and previous learning experience.

The teacher participant was a PhD student at the time of data collection, with 11 years of English teaching experience. She had worked at this research site for several years on another project but had not taught the learners prior to this study.

Instructional Treatment

There were five participant groups in this study, including a control group who only took the pretests and posttests and four experimental groups who performed two tasks under different conditions in addition to the tests. The instructional treatments were directed at the English past passive, a linguistic target that had not received any prior instruction. Out of 30 possible points, the mean pretest scores of the five groups ranged from 1.69 to 2.07 on the GJT and 1.64 to 2.10 on the EIT, indicating that the learners had no or limited knowledge about the target structure.

The four experimental groups all performed two dictogloss tasks that involved the use of the past passive and received the same length of instruction, which lasted a total of 2 hours. The FonM group performed two communicative tasks without receiving any form-focused instruction. Explicit Instruction + Task was provided with explicit instruction about the target structure before performing the two tasks. The Feedback group did not receive pretask instruction, but their wrong use of the passive voice was corrected during task performance. Finally, the Explicit Instruction + Feedback group received both pretask instruction and within-task feedback.

Dictogloss tasks. Each group performed two dictogloss tasks with identical procedures in which the students listened to a narrative presented by the teacher, worked in pairs to rehearse it, and finally were called on to retell it to the class. At the beginning of each task, the teacher went through a list of words that would appear in the narrative with the students and asked two brainstorming questions to arouse their interest and activate their schema about the topic. The teacher then read the story aloud at a normal speed to enable the students to have some initial understanding. Next, she read the story presented on PowerPoint, with one or two sentences and vocabulary annotations on each slide. Afterwards, the teacher read it a third time to solidify the students' understanding. After hearing the story three times, the students worked in pairs and practiced retelling the story with the help of clues in the form of a list of nouns and verbs. Each pair of students was asked to add an ending to the story. After the group work, members of each pair retold the story to the rest of class. At the end of each dictogloss task, the students were asked to vote on the best ending.

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

Two narrative texts were used, one for each task. One, composed by the researchers, was a news report about a car accident, and the other, which was revised based on a story from *Reader's Digest*, concerned an earthquake in Haiti. A total of 30 cases of past passives were included in the two texts. To ensure that the stories were comparable in terms of difficulty and length with the learners' school textbook, two experienced local teachers were consulted when creating the two narrative texts.

Explicit instruction. The explicit instruction in the Explicit Instruction + Task and Explicit Instruction + Feedback conditions consisted of a 15-minute mini-lesson about the meaning, form, and use of the past passive. During the lesson, the teacher started by asking the students to identify the agent and receiver of the action in the active and passive versions of the same sentence, and then explained the formation of the passive voice and the situations where it is appropriate to use the structure. Following this, the students were asked to judge the grammaticality of 10 passive sentences and provide the correct version for the incorrect sentences.

Corrective feedback. The corrective feedback consisted of corrective recasts, i.e. a prompt followed by reformulation of the wrong utterance in the absence of learner self-correction. When an error was made in using the passive voice during the oral narrative, the teacher repeated the sentence with a prosodic emphasis on the error to alert the learner to the presence of the error and encourage the learner to self-correct. If the learner succeeded in correcting the error, the teacher signaled to him or her to proceed, and no further assistance was provided. If self-correction failed, the teacher provided a recast, replacing the error with the correct form (with prosodic emphasis) while keeping the meaning intact.

Testing

A GJT and an EIT were used as respective measures of explicit and automated knowledge about the target structure before and after the instructional treatments. Both tests had three versions—pretest, immediate posttest, and delayed posttest—containing the same items but presented in different sequences. Each test included 30 items, and each correct answer received 1 point, with the total possible score being 30. Among the 30 items, 20 related to old items and 10 to new verbs that did not appear in the treatment. Three types of errors were built in the test items: (a) bare verb, as in "My dog kill in the earthquake"; (b) no past participle, as in "My dog was kill in the earthquake"; and (c) no "be," as in "My dog killed in the earthquake." The EIT was always administered before the GJT.

The GJT was untimed and required the learner to determine whether a given item was grammatically correct and, if not, to correct the error. The test included 40 items, 30 of which were target items and 10 were distractors. All target items were ungrammatical—as previous research suggests that learners' judgments of ungrammatical items provides a clearer measure of explicit knowledge (R. Ellis, 2005; Gutiérrez, 2013). The internal validity of the test (Cronbach's alpha) was .91 for the pretest and .96 for both posttests.

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

The EIT required the learner to listen to the recordings of 35 sentences read by a native speaker, decide whether each statement was true of him- or herself (e.g., "My friend was injured yesterday"), and repeat the sentence in correct English. Among the 35 items, 30 were target items and 5 were distractors. Half of the items were grammatical and half ungrammatical. Reaction time was determined by averaging the time taken to answer that item by 26 learners during a pilot test. Cronbach's alphas were .68, .76, and .77 for the pretest, the immediate posttest, and the delayed posttest, respectively.

Procedure

The study was conducted in three sessions. In session 1, the learners took the pretests; in Session 2, the experimental groups received a 2-hour treatment involving the performance of the two tasks (with a brief break between each task), followed by the immediate posttests; in Session 3, which took place 2 weeks later, the learners took the delayed posttests. In the treatment session, although the four experimental groups performed the same dictogloss tasks, the length of the tasks varied across the treatment conditions. For example, the duration of task performance was shorter in the Explicit Instruction + Task and Explicit Instruction + Feedback conditions because of the inclusion of the pretask instruction. However, as pointed out earlier, all four groups received the same length of treatment, which lasted 2 hours.

Analysis

A mixed-design repeated measure analysis was performed for the learners' scores on the GJT and the EIT to probe whether their scores varied across the treatment conditions and over time. Follow-up, one-way ANOVAs and pairwise comparisons (with Bonferroni corrections) were conducted to locate the source of any significant difference identified by the mixed-design repeated measure analysis. Nonparametric analyses were also performed due to the nonnormal distributions of some of the group scores, but similar results were obtained, so a decision was made to report the results of the parametric tests. Effect sizes (Cohen's *d*) were calculated and consulted when interpreting the results. Effect sizes were interpreted as small ($d = .40$), medium (.70), and large (1.00), following Plonsky and Oswald's (2014) recommendations. To avoid the impact of extreme values, outliers, defined as values 2.5 standard deviation units above or below the mean, were consistently excluded.

RESULTS

Overall Scores

Grammaticality judgment test. The descriptive statistics for the GJT including the means and standard deviations of each participant group's pretest and posttest

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

TABLE 1. *Means and Standard Deviations for the Grammaticality Judgment Test*

	Pretest			Posttest 1			Posttest 2		
	n^a	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Task Only	28	1.29	1.80	29	4.62	5.69	29	5.03	6.79
Explicit Instruction + Task	29	2.03	2.97	30	10.10	8.93	30	10.57	9.01
Feedback	28	1.89	3.30	30	9.43	8.19	29	6.90	8.20
Explicit Instruction + Feedback	29	2.07	2.89	29	14.97	9.63	30	12.43	11.03
Control	29	1.69	1.98	28	2.46	2.63	29	2.21	2.48

Note. [a]The number of participants ranged from 28 to 30 across groups and tests after outliers were removed.

scores are displayed in Table 1. As can be seen, the learners' scores were very low on the pretest, with means ranging from 1.29 to 2.07 out of 30, and there was very little between-group variation in the pretest scores. However, their scores, especially those of the experimental groups, increased from the pretest to the two posttests, with a slight decline from the immediate posttest to the delayed posttest in some of the groups. One observable pattern is that on both posttests, Explicit Instruction + Feedback showed the highest mean score, followed in decreasing order by Explicit Instruction + Task, Feedback, Task, and the control group.

A one-way ANOVA confirmed that there were no significant differences among the five groups' pretest scores. The mixed design ANOVA revealed a significant effect for time, $F(2, 270) = 87.51, p < .001$, for group $F(4, 135) = 9.47, p < .001$, and for group × time interaction, $F(8, 270) = 10.15, p < .001$. Given that the five groups' previous knowledge about the linguistic target did not differ significantly, any between-group differences on the posttest scores can be attributable to the differences in the type of treatment received. One-way ANOVAs and pairwise comparisons were performed on the posttest scores to locate the sources of between-group differences. At the time of the immediate posttest, all experimental groups except for Task Only significantly outperformed the control group, and all the significant differences showed large effect sizes, with the Explicit Instruction + Feedback showing the largest effect size (Table 2). At the time of the delayed posttest, the two groups receiving explicit instruction continued to score significantly higher than the control group but the other treatment groups did not. On both posttests, Explicit Instruction + Feedback also showed significantly higher scores than Task Only, with medium and large effect sizes.

Elicited imitation test. The learners' scores on the EIT appear in Table 3. Overall, the pretest scores were low (1.64–2.28), and there was limited improvement from the pretest to the posttests. There was also little between-group variation in their test scores at the three time points, despite the slightly higher posttest scores for Explicit Instruction + Feedback in comparison with the other conditions. The mixed design ANOVA only showed a significant effect for time, $F(2, 256) = 57.77, p < .001$, which is due to the overall higher posttest scores of all five

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

TABLE 2. *Post Hoc Comparisons for Treatment Effects on Grammaticality Judgment Test*

Group Contrasts	Posttest 1		Posttest 2	
	d^a	p^b	*d*	*p*
Treatment vs. Control				
Task Only vs. Control	.58	1.00	.63	1.00
Feedback vs. Control	1.12	.01*	.74	.28
Explicit Instruction + Task vs. Control	1.11	.00*	1.21	.00*
Explicit Instruction + Feedback vs. Control	1.72	.00*	1.23	.00*
Treatment vs. Other Treatments				
Explicit Instruction + Feedback vs. Task Only	1.21	.00*	.72	.01*
Explicit Instruction + Feedback vs. Explicit Instruction + Task	.52	.14	.18	1.00
Explicit Instruction + Feedback vs. Feedback	.60	.05	.55	.09
Explicit Instruction + Task vs. Task Only	.63	.06	.60	.09
Explicit Instruction + Task vs. Feedback	.06	1.00	.41	.83
Feedback vs. Task Only	.59	.15	.17	1.00

Notes. $^*p < .05$. [a]Effect size (Cohen's *d*). [b]Results of null hypothesis significance testing.

TABLE 3. *Means and Standard Deviations for Elicited Imitation Test*

	Pretest			Posttest 1			Posttest 2		
	n^a	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Task Only	29	2.28	2.28	29	3.03	2.43	30	4.00	2.85
Explicit Instruction + Task	27	1.70	2.13	29	2.69	2.93	29	3.57	3.16
Feedback	27	1.78	1.87	28	2.86	2.26	29	4.10	3.45
Explicit Instruction + Feedback	29	1.86	1.98	29	4.14	3.62	29	5.07	4.23
Control	28	1.64	1.68	30	2.53	2.36	29	2.83	2.33

Note. [a]The number of participants ranged from 28 to 30 across groups and tests after outliers were removed.

groups including the control group in comparison with their pretest scores. One-way ANOVAs failed to detect any significant differences between the five groups in terms of their pretest and posttest performance. Due to the absence of significant effects and the similarity in the group scores, post hoc analyses and effect sizes were not computed.

Zero Versus Some Previous Knowledge

Grammaticality judgment test. The second research question concerns whether the effectiveness of the instructional treatments varied as a function of whether learners had no or some previous knowledge. To address this question, a new variable was created by dividing the learners into two groups: those who scored zero on the pretest and those whose scores were above zero. The descriptive statistics for the GJT scores appear in Table 4, which shows that (a) learners with zero and some previous knowledge both improved their performance on the posttests compared with their pretest performance; (b) for both learners with and without previous

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

TABLE 4. *Means and Standard Deviations for Grammaticality Judgment Test: Zero Versus Some Previous Knowledge*

		Pretest			Posttest 1			Posttest 2		
		n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Task Only	Zero	13	0.00	0.00	13	3.23	6.66	13	4.31	8.39
	Some	15	2.40	1.84	16	5.75	4.68	16	5.63	5.38
EI[a] + Task	Zero	12	0.00	0.00	12	6.08	5.79	12	6.50	7.54
	Some	17	3.47	3.18	17	11.88	9.27	17	12.29	8.3
Feedback	Zero	12	0.00	0.00	12	4.00	4.33	11	2.82	5.02
	Some	16	3.31	3.82	17	12.12	7.41	17	8.47	8.21
EI + Feedback	Zero	11	0.00	0.00	10	9.00	8.04	11	5.64	8.15
	Some	18	3.33	3.18	18	17.50	8.89	18	15.61	10.47
Control	Zero	8	0.00	0.00	8	1.00	2.07	8	0.75	1.49
	Some	21	2.33	1.98	20	3.05	3.65	21	2.76	2.57

Note. [a]EI = explicit instruction.

TABLE 5. *Post Hoc Comparisons for Learners With Zero Versus Some Previous Knowledge on Grammaticality Judgment Test*

	Posttest 1				Posttest 2			
	Zero[a]		Some[b]		Zero		Some	
	d[c]	*p*[d]	*d*	*p*	*d*	*p*	*d*	*p*
Treatment vs. Control								
Task Only vs. Control	.41	1.00	.64	1.00	.53	1.00	.70	1.00
EI[e] + Task vs. Control	1.08	.63	1.13	.00	.96	.75	1.43	.00
Feedback vs. Control	.83	1.00	1.44	.00	.52	1.00	.83	.21
EI + Feedback vs. Control	1.29	.06	2.02	.00	.77	1.00	1.61	.00
Treatment vs. Other Treatments								
EI + Feedback vs. Task Only	.79	.23	1.48	.00	.16	1.00	1.06	.00
EI + Feedback vs. EI + Task	.42	1.00	.64	.20	− .11	1.00	.37	1.00
EI + Feedback vs. Feedback	.80	.52	.65	.25	.42	1.00	.75	.06
EI + Task vs. Task Only	.46	1.00	.67	.14	.27	1.00	.79	.12
EI + Task vs. Feedback	.41	1.00	− .01	1.00	.57	1.00	.44	1.00
Feedback vs. Task Only	.14	1.00	.88	.11	− .21	1.00	.28	1.00

Notes. [a]Learners who scored zero on the pretest. [b]Learners whose pretest scores were above zero. [c]Effect size. [d]Results of null hypothesis significance testing. [e]EI = explicit instruction.

knowledge, the group that received explicit instruction plus corrective feedback showed higher mean scores than the other groups; and (c) in all treatment conditions learners with previous knowledge showed higher gains than those without when compared with control, a pattern which can be more clearly observed by comparing the effect sizes for the "zero" and "some" learners in each group in Table 5.

The mixed design ANOVA for the GJT for the learners with no previous knowledge only detected a marginally significant main effect for time, $F(4, 50) = 2.55$, $p = .05$. Similarly, one-way ANOVAs did not show any significant between-group differences in the pretest or posttest scores. However, as Table 5 shows, despite the nonsignificant p values, the experimental groups, except for Task Only, all showed

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

TABLE 6. *Means and Standard Deviations for Elicited Imitation Test: Zero Versus Some Previous Knowledge*

		Pretest			Posttest 1			Posttest 2		
		n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Task Only	Zero[a]	9	0.00	0.00	9	1.78	1.99	9	2.22	1.99
	Some[b]	20	3.3	2.03	19	3.79	2.35	20	4.95	2.80
EI[c] + Task	Zero	11	0.00	0.00	11	0.73	1.27	10	1.30	1.25
	Some	16	2.88	2.06	17	4.12	2.96	17	5.12	3.06
Feedback	Zero	8	0.00	0.00	8	1.25	1.75	7	1.71	3.40
	Some	19	2.53	1.74	18	3.67	2.09	20	5.20	3.09
EI + Feedback	Zero	10	0.00	0.00	9	1.22	0.67	10	1.10	1.66
	Some	19	2.84	1.77	20	5.45	3.65	19	7.16	3.62
Control	Zero	9	0.00	0.00	9	1.22	2.44	9	1.11	1.45
	Some	19	2.42	1.50	19	3.11	2.16	18	3.61	2.38

Notes. [a]Learners who scored zero on the pretest. [b]Learners whose pretest scores were above zero. [c]EI = explicit instruction.

large effect sizes on Posttest 1 and Explicit Instruction + Feedback continued to show a large effect at the time of Posttest 2, suggesting that the nonsignificant p values may have been due to the smaller sample sizes for the zero groups ($n = 8 - 13$).

The ANOVA results for the GJT scores for the learners with some previous knowledge are similar to the patterns for the overall GJT scores (for the whole cohort). The conditions with explicit instruction and/or feedback showed superior immediate effects compared with control, and the two conditions with explicit instruction continued to show significantly higher scores than control at the time of the delayed posttest. Furthermore, Explicit Instruction + Feedback significantly outperformed Task Only on the immediate posttest, and this condition also showed consistently larger effect sizes than other treatment types.

Elicited imitation test. The descriptive statistics displayed in Table 6 reveal that for learners with zero previous knowledge, there was little between-group variation in their posttest scores, and there was also limited within-group improvement from the pretest to the posttests. The mixed design ANOVA only revealed a significant main effect for time, $F(2, 78) = 17.65$, $p < .001$, which was due to the overall higher posttest scores in comparison with the pretest scores (which were zero in this case). However, there were no significant effects for group or group × time interaction. One-way ANOVAs did not detect any differences among the groups in their pretest and posttest performance, and the effect sizes (Table 7) for the treatment groups compared with the control group were either small or negative (except for a medium effect size for Task Only).

Unlike the results for learners without prior knowledge or for overall EIT scores, the results for the learners with prior knowledge not only showed a significant effect for time, $F(2, 168) = 43.33$, $p < .001$, but also for group, $F(4, 84) = 17.65$, $p = .03$, and time × group interaction, $F(8, 168) = 17.65$, $p < .001$. One-way ANOVAs showed that there were no significant differences in the groups' pretest scores or

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

TABLE 7. *Post Hoc Comparisons for Learners with Zero Versus Some Previous Knowledge on Elicited Imitation Test*

	Posttest 1				Posttest 2			
	Zero[a]		Some[b]		Zero		Some	
	d[c]	*p*[d]	*d*	*p*	*d*	*p*	*d*	*p*
Task Only vs. Control	.25	1.00	− .09	1.00	.64	1.00	.18	1.00
Explicit Instruction + Task vs. Control	− .26	1.00	.22	1.00	.14	1.00	.38	1.00
Feedback vs. Control	.01	1.00	.21	1.00	.24	1.00	.53	1.00
Explicit Instruction + Feedback vs. Control	.00	1.00	.64	.09	− .01	1.00	1.02	1.00

Notes. Due to lack of significant effects and small mean differences between the treatment groups, effect sizes were only calculated for treatment-control contrasts, not for contrasts between the treatment groups. [a]Learners who scored zero on the pretest. [b]Learners whose pretest scores were above zero. [c]Effect size. [d]Results of null hypothesis significance testing.

their scores on the immediate posttest, but there were in their scores on the delayed posttest, $F(4, 89) = 3.27$, $p = .02$. Post hoc Bonferroni tests found a significant superior effect for Explicit Instruction + Feedback in comparison with control, with a large effect size, $d = 1.02$.[1]

DISCUSSION

The study was designed to investigate the effect on learning of three different approaches for using tasks in language pedagogy. In FonM the learners performed two dictogloss tasks without either explicit instruction or corrective feedback. In FonF, the learners performed the tasks and received corrective recasts when they committed errors in the target structure (past passive). This corresponds to what Long (2015) would consider "pure task-based teaching." There were also two conditions involving TSLT. The learners received explicit instruction in one of these conditions before performing the task and explicit instruction together with corrective feedback while performing the task in the other. We examined what effects these different approaches had on the acquisition of both explicit knowledge, measured by a GJT, and on more automated knowledge, measured by an EIT. We also investigated whether there was any difference in these effects for learners who had no prior knowledge of past passive and those who possessed some limited knowledge.

FonM (Task Performance With No Requirement for Any FonF)

In some studies involving dictogloss tasks learners engage extensively in what Swain (2006) called "languaging" (i.e., talking explicitly about the choice of linguistic forms). However, such studies have typically involved advanced, adult learners. In our study, the learners were children, aged 14 or more years old, with limited language proficiency. In our study, too, the learners were only asked to

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

reproduce the narratives orally (i.e., they did not write them). An inspection of the interactions that took place in the learner groups and in the whole-class oral presentation of the narratives indicated that almost no "languaging" took place. The learners focused on entirely on reproducing the content. Thus we consider that the dictogloss tasks effectively led to FonM.

However, the dictogloss tasks did provide learners with input containing exemplars of past passive but did not make its use essential and, in fact, the learners made few attempts at producing it. Thus the FonM condition provided some opportunity for learning the target structure but this was clearly limited. In fact, the FonM condition resulted in very little learning. In the GJT small increases from pretest to both posttests occurred but these were not significantly greater than those observed for the control group and the effect sizes for the Task Only or control group comparisons were small ($d = .58$ and .63). The gains from pre- to posttests in the EIT were even smaller and scarcely greater than those for the control group. This absence of any effect for FonM was evident for both learners with some prior knowledge of past passive and those with no knowledge.

There are three possible explanations for this failure. First, past passive is a late acquired feature, and the learners were not developmentally ready. However, as we will see, some learning of this structure was evident in the other groups, especially in those learners with some prior knowledge and who therefore may have been more developmentally ready. Second, performing just two tasks did not provide sufficient opportunities for incidental learning to take place. This seems likely given that such learning is a slow, dynamic process that requires extensive and ongoing experience with the language and the task provided very limited experience of the target structure. Third, as Long and Robinson (1998) pointed out, when there is an exclusive FonM—which, as we argue, occurred in this condition—the learners probably failed to activate the cognitive processes required for acquisition to take place.

FonF (Task-Based Language Instruction)

In the FonF condition, the learners potentially benefitted from the exposure to the exemplars of past passive in the input provided by the dictogloss tasks, while, in addition, the corrective recasts they received while performing the tasks may have helped to attract attention to the use of the past passive and also to encourage the learners to notice the gap between their own erroneous utterances and the target form and thus engage in the cognitive comparison in the window of opportunity that has been hypothesized to promote acquisition (Doughty, 2001). Doughty and Varela (1998) considered corrective recasts a relatively implicit form of corrective feedback. However, their frequency and the fact that the learners' erroneous utterances were repeated before they were recast likely made the feedback salient to the learners.

The results show a clear effect for the FonF treatment in the acquisition of explicit knowledge. In the immediate GJT posttest the FonF group outperformed the control group with a large effect size ($d = 1.12$). However, this effect decayed over time. By the delayed posttest, the difference between the FonF group and the

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

control group was no longer significant, and the effect size was small ($d = .63$). Small increases in the EIT were also evident and, interestingly, continued from the immediate to the delayed posttest. However, there were no statistically significant differences between the FonF and control groups in either of the posttests and the effect sizes were negligible. As Table 6 shows, in comparison with the control group, those learners with some prior knowledge of the target structure developed the highest level of automated knowledge in the delayed posttest, but the effect size ($d = 53$) was small.

These results suggest that the effects of the FonF treatment were largely limited to explicit knowledge. This is not surprising. Studies (e.g., Morgan-Short et al., 2015) have shown that the initial stage of learning is explicit. The short length of the instruction in this study seems to have helped the learners to form an explicit representation of this new target structure but was insufficient to develop automated knowledge. The results also support the claim of Berry and Dienes (1993) and Reber (1989) that explicit knowledge decays more rapidly than implicit knowledge. As explicit knowledge decays, automated knowledge develops; this seems to have been the case for those learners who had some prior knowledge of the target structure. More extensive FonF instruction may have resulted in clearer evidence of the development of automated knowledge.

Explicit Instruction + Task (Task-Supported Language Teaching)

It should be noted that this treatment did not include the extensive controlled practice activities that, for example, Ur (1996) proposed for teaching new grammar points and thus does not constitute a full version of PPP. However, it enables us to examine whether equipping learners with explicit information about past passive resulted in more effective use of the learning opportunities provided by the dictogloss tasks. Our principal interest, then, is whether this treatment led to greater learning than the FonM condition (i.e., task alone).

The GJT results demonstrate that when these learners were provided with explicit instruction, they outperformed the control group ($p < .05$) in both posttests with large effect sizes ($d = 1.11$ and 1.21). The comparison with the Task Only (FonM) group approached statistical significance ($p = .06$ in Posttest 1 and .09 in Posttest 2) but with small effect sizes ($d = .63$ and .60). Notably, and in contrast to the FonF group, GJT scores for this group did not decline between the two posttests. Those learners with some existing knowledge of past passive performed more strongly on the GJT than those with no knowledge, and in fact, it was only the former that demonstrated significant differences compared to the control group. The results for the EIT show that providing learners with explicit information before they performed the task contributed little to the development of automated knowledge. Scores did increase for this group across the three tests, most clearly for those learners with some prior knowledge of the target structure. However, none of the comparisons between the Explicit Instruction + Task group and either

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

the control group or the FonM group reached statistical significance and effect sizes were negligible or small.

Clearly, there was some benefit in providing learners with explicit knowledge before the task, but this was restricted to the effect it had on learners' explicit knowledge, which was at least durable (i.e., it did not decline as occurred in the FonF group). There was no clear effect on their automated knowledge. As noted in the introductory sections of this chapter, DeKeyser (1998) argued that "proceduralization is achieved by engaging in the target behavior—or procedure—while temporarily leaning on declarative crutches" (p. 49). In the context of this study, this means that the learners needed to draw on the declarative knowledge they gained from the explicit instruction while they performed the task. But the task was challenging and the learners overall L2 proficiency was also limited, which must have made it difficult for them to move beyond the primary FonM the tasks required. It is unlikely, then, that and they were able to "lean on the declarative crutches" (p. 49) that the explicit instruction had afforded them as they performed the task.

Explicit Instruction + Task + Corrective Feedback (Stronger Version of Task-Supported Language Instruction)

This condition differs from the former task-supported condition because it also incorporated corrective feedback. In effect, this condition also included FonF via the corrective feedback that arose incidentally while learners were primarily engaged with communicating. As R. Ellis (2015) noted, PPP can also include FonF if it incorporates a free production task. This condition, however, did not include any controlled practice activities of the kind Ur (1996) described.

This treatment produced the strongest effect on explicit knowledge. GJT scores increased substantially from pretest to the immediate posttest, although, like the FonF group, they declined from the immediate to the delayed posttest. This group outperformed the control group with large effect sizes at both posttest times ($d =$ 1.72 and 1.23) and also the FonM group ($d = 1.21$ and .72). It also outperformed the FonF group on the posttests, although statistical significance was only achieved for the immediate posttest and the effect size was small ($d = .60$).

This treatment also had the strongest effect on automated knowledge, but for the whole sample none of the differences with the other groups (including the control group) reached statistical significance. However, when just the learners with prior knowledge were considered, this group achieved higher EIT scores than all the other groups, and the comparison with the control group in the delayed posttest reached statistical significance with a large effect size ($d = 1.02$).

The corrective feedback that the learners received seems to have helped to consolidate the explicit knowledge of the target structure provided through the explicit instruction. The feedback provided them with examples of the rules they had been taught so they could make links between concrete exemplars of past passive available while they performed the tasks and the abstract rules provided by the explicit instruction. It would seem, though, that the explicit representations they formed were subject to decay. Perhaps, they faded as procedural memory took

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

over and automated knowledge gained ground. If this was what happened, then the pattern of results is again best explained by a theoretical model that views the initial stage of the learning of a new feature as explicit in nature, which is replaced over time as procedural learning takes place and automated knowledge develops. What is clear is that learners who start off with some declarative knowledge of a "new" feature are better equipped to benefit from the combination of explicit instruction and corrective feedback than those with no such knowledge.

Summary

Overall, those treatments that involved attention to the form of the target structure (i.e., the two task-supported treatments and FonF) were more effective than the two conditions (FonM and control) where there was no such attention. Furthermore, the more explicit the treatment, the greater the effect. That is, the task-supported group that received explicit instruction followed by within-task feedback showed a consistent advantage over the other treatment conditions. However, these effects were largely limited to the acquisition of explicit knowledge as none of the groups outperformed the control group on the EIT. In general, then, the treatments were not successful developing the learners' automated knowledge.

The treatments were clearly more effective for those learners who started with some previous explicit knowledge. However, those learners with no prior explicit knowledge also benefitted, as shown by the medium to large effect sizes for the GJT comparisons between the FonF and the two task-supported groups and the control group. The advantage for those learners with some prior knowledge was more clearly evident in the EIT scores. The learners with prior knowledge in all four experimental groups outscored the learners with zero knowledge, especially in the group that received both explicit instruction and corrective feedback. In short, prior knowledge together with the most explicit type of instruction worked best for the development of automated knowledge (but see endnote 1).

These results are clearly compatible with a skill-learning theoretical model (DeKeyser, 1998, 2007) that proposes automated knowledge develops when learners have access to an explicit representation of a grammatical feature and the opportunity to utilize this knowledge while engaging in communicative behavior where their attention is repeatedly drawn to the target feature. Such a model views explicit knowledge as transforming into implicit knowledge through practice. But the results are also compatible with R. Ellis's (1994) and N. Ellis's (2005) contention that explicit knowledge does not transform into implicit knowledge but rather facilitates the statistical learning processes involved in its development. Thus the study cannot be seen as a test of the strong- and weak-interface positions.

CONCLUSION

The results of this classroom-based study suggest that TSLT rather than "pure" TBLT holds more promise for teaching grammatical structures that are new. However, some caveats are in order.

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

First, the learners in the study were young adolescents who were well-attuned to form-focused instruction of the explicit kind and thus, perhaps, naturally responsive to receiving explicit information about a grammatical structure. Younger learners or even older learners who are more functionally oriented to learning a L2 may benefit more from "pure" TBLT.

Second, there is ample evidence to suggest that TBLT directed at grammatical structures that are already partially acquired is effective in pushing learners through an acquisition sequence (e.g., Mackey, 1999) and in improving accuracy (e.g., Doughty & Varela, 1998; Yang & Lyster, 2010). Thus, we do not wish to suggest that TSLT is preferable for such structures. Further research is needed to investigate this. We note, however, that the bulk of the task-based research to date has focused on partially acquired structures and suggest that more research is needed that focuses on new structures.

Third, the experimental nature of our study, where the instruction was limited to just 2 hours, is arguably biased in favor of explicit instruction. Two hours is very little for any measurable effect for the incidental learning of a difficult grammatical structure to become apparent. It is possible that different results would be obtained in a longitudinal study where the gradual and accumulative effect of FonF is more likely to become evident. Enough is now known about the complex nature of the acquisition of L2 grammar to discount simplistic claims about which approach—task-based or task-supported—is superior.

Fourth, while it is clearly desirable to investigate what effect TBLT has on the acquisition of specific grammatical structures—as in our study—we also recognize that the goals of TBLT are not restricted to the acquisition of grammar. TBLT seeks to develop the all-around, holistic competence of the L2 learner, and our study did not investigate the effect of the different conditions on the development of such a competence.

Fifth, there are a number of methodological issues we would like to draw attention to. The groups receiving explicit instruction received more exposure to the target constructions. Also, because we controlled for time in all the conditions, the explicit instruction groups spent less time on task-based work. Perhaps, the effect of the explicit instruction would have been greater if these groups had spent the same amount of time on performing the tasks as the other two groups.

Finally, we acknowledge that this study was entirely product-based. We have not reported what effect the different instructional conditions had on the learners' production of the target feature as they performed the tasks or, crucially, on whether there was a differential effect on fluency or overall accuracy. We plan to investigate these issues in a forthcoming study.

NOTE

1 Further analysis showed that the significant difference was only evident for grammatical items, not for ungrammatical items, suggesting that the learners had developed a degree of proceduralization that enabled them to reproduce past passive verb forms correctly but had not achieved the fully automated knowledge of past passive required to correct erroneous forms automatically.

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

REFERENCES

Andringa, S., de Glopper, K., & Hacquebord, H. (2011). Effect of explicit and implicit instruction on free written response task performance. *Language Learning*, *61*(3), 868–903.

Basturkmen, H., Loewen, S., & Ellis, R. (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Applied Linguistics, 25*(2), 243–272.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hove, UK: LEA.

Bowles, M. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute? *Studies in Second Language Acquisition*, *33*(2), 247–271.

Cintrón-Valentín, M., & Ellis, N. (2015). Exploring the interface: Explicit focus-on-form instruction and learned attentional biases in L2 Latin. *Studies in Second Language Acquisition. 37*(2), 197–235.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). Cambridge, UK: Cambridge University Press.

DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition*. Malden, MA: Blackwell.

DeKeyser, R. (2007). Skill acquisition theory. In J. Williams & B. VanPatten (Eds.), *Theories in second language acquisition: An introduction* (pp. 97–113). Mahwah, NJ: Erlbaum.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). Cambridge, UK: Cambridge University Press.

Doughty, C., & Varela, E. (1998). Communicative focus-on-form. In C. Doughty & J. Williams (Eds.), *Focus-on-form in classroom second language acquisition* (pp. 114–138). Cambridge, UK: Cambridge University Press.

Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, *27*(2), 305–352.

Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 79–114). San Diego, CA: Academic Press.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*(2), 141–172.

Ellis, R. (2006). Researching the effects of form-focused instruction on L2 acquisition. *AILA Review*, *19*, 18–41.

Ellis, R. (2007). Explicit form-focused instruction and second language acquisition. In B. Spolsky & F. Hult (Eds.), *The handbook of educational linguistics* (pp. 437–455). Oxford, UK: Blackwell.

Ellis, R. (2015). *Understanding second language acquisition*. Oxford, UK: Oxford University Press.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Pre-emptive focus on form in the ESL classroom. *TESOL Quarterly, 35*(3), 407–432.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, *28*(2), 339–368.

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition*. London, UK: Routledge.

Goo, J., & Mackey, A. (2013). The case against the case for recasts. *Studies in Second Language Acquisition*, *35*(1), 127–165.

Gutiérrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition*, *35*(3), 423–449.

Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford, UK: Oxford University Press.

Hulstijn, J. H., Ellis, R., & Eskildsen, S. W. (Eds.). (2015). Orders and sequences in the acquisition of L2 morphosyntax, 40 years on [Special issue]. *Language Learning, 65*(1).

Li, S. (2010). The effectiveness of corrective feedback in SLA: a meta-analysis. *Language Learning, 60*(2), 309–365.

Li, S., Zhu, Y., & Ellis, R. (in press). The effects of the timing of corrective feedback on the acquisition of a new linguistic structure. *Modern Language Journal*.

Lightbown, P. (2008). Transfer appropriate processing as a model for classroom second language acquisition. In Z. Han (Ed.), *Understanding second language process* (pp. 27–44). Clevedon, UK: Multilingual Matters.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. *In K.* de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam, The Netherlands: John Benjamins.

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition*. San Diego: Academic Press.
Long, M. (2015). *Second language acquisition and task-based teaching*. Malden, MA: Wiley-Blackwell.
Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition* (pp. 15–41). Cambridge, UK: Cambridge University Press.
Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 123–167). Clevedon, UK: Multilingual Matters.
Lyster, R. (2001). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning, 51*, 265–301.
Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition, 26*(3), 399–432.
Lyster, R., & Ranta, L. (2013). The case for variety in corrective feedback research. *Studies in Second Language Acquisition*, *35*(1), 167–184.
Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA. *Studies in Second Language Acquisition, 32*, 265–302.
Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, *21*, 557–587.
Morgan-Short, K., Deng, Z., Brill-Shuetz, K., Faretta-Studcenberg, M., Wong, P., & Wong, F. (2015). A view of the neural representation of second language syntax through artificial language learning under implicit and explicit contexts of exposure. *Studies in Second Language Acquisition*, *37*(2), 383–419.
Muronoi, H. (2000). Focus on form through interaction enhancement. *Language Learning*, *50*(4), 617–673.
Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417–528.
Pienemann, M. (2005). An introduction to processability theory. In M. Pienemann (Ed.), *Cross-linguistic aspects of processability theory* (pp. 1–60). Amsterdam, The Netherlands: John Benjamins.
Plonsky, L., & Oswald, F. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, *64*(4), 878–912.
Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219–235.
Reber, P., & Squire, L. (1998). Encapsulation of implicit and explicit memory in sequence learning. *Journal of Cognitive Neuroscience*, *10*(2), 248–263.
Sanchez, D., & Reber, P. (2013). Explicit pre-training instruction does not improve implicit perceptual-motor sequence learning. *Cognition*, *126*(3), 341–351.
Sanz, C., & Morgan-Short, K. (2004). Positive evidence vs. explicit rule presentation and explicit negative feedback: A computer-assisted study. *Language Learning, 54*(1), 35–78
Sato, R. (2010). Reconsidering the suitability and effectiveness of PPP and TBLT in the Japanese EFL classroom. *JALT Journal, 32*(2), 189–200.
Shanks, D. R. (2005). Implicit learning. In K. Lamberts & R. Goldstone (Eds.), *Handbook of cognition* (pp. 202–220). London, UK: Sage.
Shehadeh, A. (2005).Task-based language learning and teaching: Theories and applications. In C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in English language teaching* (pp. 13–30). New York, NY: Palgrave Macmillan.
Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(1), 38–62.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.
Spada, N., Jessop, L., Suzuki, W., Tomita, Y., & Valeo, (2014). Isolated and integrated form-focused instruction: Effects on different types of L2 knowledge. *Language Teaching Research, 18*, 453–473.
Spada, N., & Tomita, . (2010). *Interactions between type of instruction and type of language feature: A meta-analysis. Language Learning*, *60*(2), 263–308.
Suzuki, Y., & DeKeyser, R. (2015). Does elicited imitation measure implicit knowledge? Evidence from the word-monitoring task. *Language Learning*, *65*, 860–895.
Swain, M. (2006). Languaging, agency and collaboration in advanced second language learning. In H. Byrnes (Ed.), *Advanced language learning: The contributions of Halliday and Vygotsky* (pp. 95–108). London, UK: Continuum.

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press

Ur, P. (1996). *A course in language teaching: Practice and theory*. Cambridge, UK: Cambridge University Press.

Williams, J., & Evans, J. 1998. What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus-on-form in classroom second language acquisition* (pp. 139–155). Cambridge, UK: Cambridge University Press.

Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners? Acquisition or regular and irregular past tense forms. *Studies in Second Language Acquisition, 32*(2), 235–263.

Zhang, Y. (2015). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition, 37*, 457–486.

https://doi.org/10.1017/S0267190515000069 Published online by Cambridge University Press