9

Computerized Type of Feedback and Depth of Processing During a Computerized Problem-Solving Task

Hui-Chen Hsieh

Introduction

The computer has been employed in several studies to provide some type of feedback to second/foreign language (L2) learners in an effort to promote L2 development (e.g., Cerezo, Caras, & Leow, 2016; Rosa & Leow, 2004). For decades, using the computer to assist language instruction or learning has been promoted in L2 classrooms as a push from generational arguments such as the net generation of digital natives (Prensky, 2001) or from second language acquisition (SLA) pioneer research study (Chapelle, 2003). However, decision making should include considering the findings of theoretically driven and empirically supported research (cf. Cerezo, Baralt, Suh, & Leow, 2014; Jones & Shao, 2011). Thus, a careful analysis of features that the computer provides in the L2 classroom, such as feedback, is indispensable.

Feedback has been a focus in Instructed SLA (ISLA) over the last 35 years (e.g., Cerezo, 2016; Nagata, 1993). Yet, as to its effectiveness on L2 development, the jury is still out since feedback has been confounded with other intervening variables or provided at different stages along the L2 learning process. At the same time, *how* deeply such feedback has been processed by the L2 learner has not been fully investigated. To this end, the present study proposes a re-analysis of think-aloud protocols verbalized while processing feedback to address this issue.

Review of the Literature

Feedback

Feedback, commonly referred to as the information provided to confirm or revise learners' interlanguage (Loschky & Bley-Vroman, 1993) or to notice the disparity between learner's interlanguage and target language (e.g., Robinson, Mackey, Gass, & Schmidt, 2012), has been investigated over the last 25 years in computer-assisted language instruction (CALI) or non-CALI strands. Several meta-analyses (e.g., Li, 2010; Mackey & Goo, 2007) have reported overall beneficial effects of feedback during interaction on L2 development while underscoring other potential mediating factors (e.g., individual differences in cognitive capacities, L2 proficiency

level, prior knowledge, complexity of linguistic structure, age, or different settings) that may moderate the positive effects of feedback on L2 development (e.g. Yilmaz, 2012).

In the CALI strand of research (e.g., Bowles, 2008; Cerezo et al., 2016; Diez-Bedmar & Perez-Paredes, 2012; Loewen & Erlam, 2006; Nagata, 1993; Nagata & Swisher, 1995; Rosa & Leow, 2004; Sanz & Morgan-Short, 2004; Yun, Miller, Baek, Jung, & Ko, 2008), many published studies have focused on the efficacy of feedback with different degrees of explicitness to draw a learner's attention to (or awareness of) the target items in order to trigger certain L2 development. Explicit feedback has mainly been operationalized as a piece of metalinguistic explanation provided to notify learner accuracy, or lack thereof, of a response, whereas less explicit (implicit) feedback simply informs learners whether an answer is right or wrong or provides a prompt to process the feedback a bit deeper. Generally, findings appear to support positive effects of feedback on L2 development, but the effectiveness of different types of feedback is less agreed upon (cf. e.g., Bowles, 2008; Nagata, 1993; Nagata & Swisher, 1995; Rosa & Leow, 2004; Yun et al., 2008 for positive effects for explicit feedback and Loewen & Erlam, 2006; Rosa & Leow, 2004; Sanz & Morgan-Short, 2004 for no effect for type of feedback). A close examination of these studies revealed that the learning task, prior knowledge, linguistic complexity, and the learning processes may all contribute in some way to the difficulty of making a direct comparison among these studies. Moreover, the explicitness or implicitness of feedback provided to learners is simply a methodological manipulation from the researcher's perspective. However, from the learner's perspective, the perception of feedback as explicit or implicit may potentially be different. At the same time, given that learners are the main processors of L2 information, investigating how they process such feedback clearly needs to be documented in order to explicate the findings.

One recent study, Cerezo (2016), addressed whether a learning task that includes differences between the amount (28 vs. 56 items) and type of input-based practice (two vs. four options), as postulated by Sanz and Morgan-Short (2004), could explain the conflicting results in those studies (Rosa & Leow, 2004 vs. Sanz & Morgan-Short, 2004) that compared different types of feedback. Cerezo reported no evidence to support either postulations but also reported interesting data to account for the non-significant benefits of the 56-item group (extra practice) over the 28-item group: Concurrent data (think-aloud protocols) revealed several participants complaining about the length of the treatment leading to boredom and/or fatigue. The potential roles of variables such as boredom, fatigue, prior knowledge, complexity of target form or structure and so on, and how learners process the L2 information clearly need to be established during treatment phases of a study in an effort to shed light on what takes place during the learning process.

Furthermore, it is also possible that the effects of feedback may rest on its potential for leading learners to perceive its corrective function and drawing learners' attention to relevant linguistic input that triggers language encoding, decoding, or restructuring depending on which learning stage the learner is at. Although some studies have measured learners' online noticing (e.g., Bowles, 2008; Rosa & Leow, 2004), none has investigated the relationship between learner's noticing and type of feedback. Interestingly, Hsieh (2007) sought to tease out the features of task-essentialness (Loschky & Bley-Vroman, 1993) by investigating the effects of input/receptive practice (cf. Leow, 2007) and type of feedback premised on learners' attention and/or awareness. This study exposed first-semester college-level adult learners to the complex Spanish *gustar* (to like) under six different conditions where participants were required to actively select ([+practice]) the L2 input data to process or to observe (exposure-to-practice or [-practice]) the process of L2 input data being selected together with the provision of implicit feedback, explicit feedback, or no feedback, respectively. The statistical analyses revealed that computerized practice did not seem to play a significant role in learners' development of the verb as measured by

oral, written, and recognition tests given the medium or low effect sizes and power. In contrast, computerized feedback, implicit or explicit, during practice appeared to significantly promote *gustar* development given the large partial eta squared and observed power values obtained on all three types of post-exposure performances (see Appendix A). However, with respect to the effects of type of feedback, the two computerized feedback groups were not significantly different from each other on all three post-exposure tests. It is important to analyze the online attentional data collected and then to explore how learners processed feedback. To this end, a brief introduction of the recent model proposed by Leow (2015) may provide a framework to identify a learner's inner resources while processing written-online-feedback information in order to potentially address the existing gaps among computerized feedback studies.

Model of the L2 Learning Process in ISLA (Leow, 2015)

Premised on the role of attention in the process of learning an L2 and based on empirical data collected for the last two decades, Leow (2015) proposed a new model that comprises five stages of two products and three processes that may be influenced by other inner resources to exhibit the L2 learning process in instructed SLA. Of the influential inner resources, depth or level of processing is proposed to provide insights into instances of L2 learning or lack thereof to explicate the inconsistent findings among experimental groups in different strands of research in ISLA (Leow, 2015). Of interest in this chapter is the intake processing stage, in which depth of processing is postulated to play an important role in the learning process. In other words, *how* L2 learners interact with feedback should shed light on any subsequent potential learning of the target items in the input.

DoP is defined as

> the relative amount of cognitive effort, level of analysis, and elaboration of intake, together with the usage of prior knowledge, hypothesis testing, and rule formation employed in decoding and encoding same (sic) grammatical or lexical item in the input.
>
> *(Leow, 2015, p. 204)*

In his review of DoP, Leow (2015) traced an early reference to levels of processing in the field of cognitive psychology in the 1970s (Craik & Lockhart, 1972), but has not been widely investigated in the field of (I)SLA with respect to grammatical information until recently (Hsieh, Moreno, & Leow, 2016; Leow, Hsieh, & Moreno, 2008; Morgan-Short, Heil, Botero-Moriarty, & Ebert, 2012; Qi & Lapkin, 2001; see Hulstijn, 1997; Laufer & Hulstijn, 2001; Rott, 2005 for lexical learning). Results from these empirical studies generally point to a positive role of DoP in L2 development. However, these findings are generally reported with some limitations, such as low number of participants, lack of online measure, and so on. To further explore this line of research, Adrada-Rafael (2017) recently exposed 88 learners to Spanish imperfect subjective in a reading passage under three different experimental conditions (rule explained explicitly, rule searching required-implicit, and baseline) and found that the more explicit condition yielded more instances of deeper processing and a better performance at controlled production and reading comprehension.

Research Question

To expand this line of research to give a more complete picture to the L2 research community, the present study proposed to revisit Hsieh (2007) to analyze the online attentional data in

relation to feedback by following Leow's (2015) recent model to probe whether DoP did play a role in the intake processing stage. To this end, the research question was as follows:

> Does type of computerized feedback (explicit vs. implicit vs. no feedback) provided to adult learners while processing Spanish *gustar* in a computerized puzzle game elicit differential levels of processing?

Research Design

Participants

The original participant pool contained 110 college-level students of Spanish enrolled in a first-semester Spanish level, randomly assigned to six treatment conditions, which resulted from the combination of [±practice] and [±feedback] ([+feedback] was further divided into [+implicit feedback] and [+explicit feedback]). Forty-seven participants were eliminated due to high pretest scores, absence or withdrawal from the course, or incomplete data due to technical problems. The remaining 63 participants (29 females, 34 males, and ages ranging from 17 to 34 with an average age of 19.2 years) were included in the participant pool for statistical analysis. The participants met three times a week; classes were conducted almost exclusively in Spanish; and the development of the four language skills was focused in the course.

The cut-off point at the pretest was 40% given that the *gustar* structure is well-known even among beginner learners who have been exposed to it during high school.

The Target Structure

The verb *gustar* 'to like' typically contains an animate entity as the indirect object and a thing/an object as the subject. This structure is difficult to learn due to a combination of factors. Semantically, the meaning of *gustar* is 'to be pleasing' or 'to please,' but learners tend to interpret it as the English literary equivalent 'to like.' Compare (1) and (2a, b) below.

1. I like the movie.
   I-nom. like the movie-acc.
2. a. A mí me gusta la película.
      To me-dat. 1st-dat. pleases the movie-nom.
   b. La película me gusta a mí.
      The movie-nom. 1st-dat. pleases to me-dat.

Syntactically, (1) 'to like' is a transitive verb that takes an accusative case while the intransitive verb *gustar* does not (Whitley, 1995); (2) 'to like' acts as an agentive verb and follows canonical (SVO) word order, whereas *gustar* requires a dative experiencer and does not necessarily follow the canonical word order. Normally, learners use the first noun strategy (NounVerbNoun = SubjectVerbObject) to process Spanish and overlook the morphosyntactic feature of the input (Montrul, 1997; Sanz, 1999; VanPatten, 1996). Thus 'I like it' is mapped onto '*me gusta,*' in which '*me*' is wrongly interpreted as the subject 'I' and the verbal morphology is, then, filtered out. Subsequently, other erroneous expressions are derived as '**(Yo)*

me gusto las manzanas' or '**Yo me gusta/gustan las manzanas'* instead of '(A mí) *Me gustan las manzanas*.'

Gustar, morphologically, requires obligatory clitic-doubling with dative experiencers and the case marker *a*, which are unstressed, monosyllabic, and quite difficult to be perceived. In addition, the clitic marks for case, number, and gender simultaneously, which make it quite challenging to learn (Sanz, 1999). Consequently, learners normally produce expressions like '**A Juan les/los gustan los deportes*' or '**A Juan las gustan las casas*' in which they mismatch the indirect object pronoun with the nominative case in number and gender.

Types of Gustar *Structure Used in the Treatment*

There were 12 expressions of *gustar* according to the different combinations of the subject and the indirect object. In order to reduce the complications, only the expressions of *yo*, *él*/*ella*, and *ellos*/*ellas* were included in the study. Every expression was repeated four times to balance and match the gender and the number of the subject and indirect object in order to help participants reject the erroneous hypotheses. Consequently, there were a total of 24 target sentences evenly distributed among the following structures:[1]

a. Me gusta. . . /Me gustan. . .
b. Le gusta. . . /Le gustan. . .
c. Les gusta. . . /Les gustan. . .

The Treatment Task

A problem-solving task with features like those found in many computerized games such as Pac-Man (cf. Bowles, 2008 for the use of such a problem-solving task to draw learners' attention to targeted items in the input) was created via *Macromedia Flash* including various flash cards with correct linguistic items of *gustar*, distractors, metalinguistic explanation as explicit feedback and a wall as the implicit feedback to describe the pictures using the *gustar* expressions (see Figure 9.1).

The participants in the [+practice] groups had to select any path(s) to form a sentence to describe the picture. Along the pathway they selected, either no feedback, implicit feedback, or explicit feedback was provided to confirm or disconfirm their choice of the path(s). Meanwhile, participants in the [-practice] (the exposure to the [+practice] or "exposure-to-practice") groups observed their partner in the [+practice] group via a monitor that was paired up by using *Screen sharing*. Table 9.1 summarizes the experimental conditions of the present study.

The experimental condition PRONLY allowed its participants to move down the path to form a complete sentence they thought would describe the picture with no feedback provided; while the participants in PRIFB could continue unhindered if the choice was correct, or, if an incorrect choice was made, a wall appeared to block them from moving forward unless they went back to pick the correct expression. Finally, PREFB exposed the participants to the same material as in the PRIFB and an additional explicit metalinguistic explanation for every path they selected.

Synchronized to every participant in these [+practice] groups, there was one in the [-practice] groups who had no control over what linguistic input to select along the path, but sat in front of a computer monitor to simultaneously observe (be exposed to) what their partners selected, viewed, and read, and the type of feedback they received.

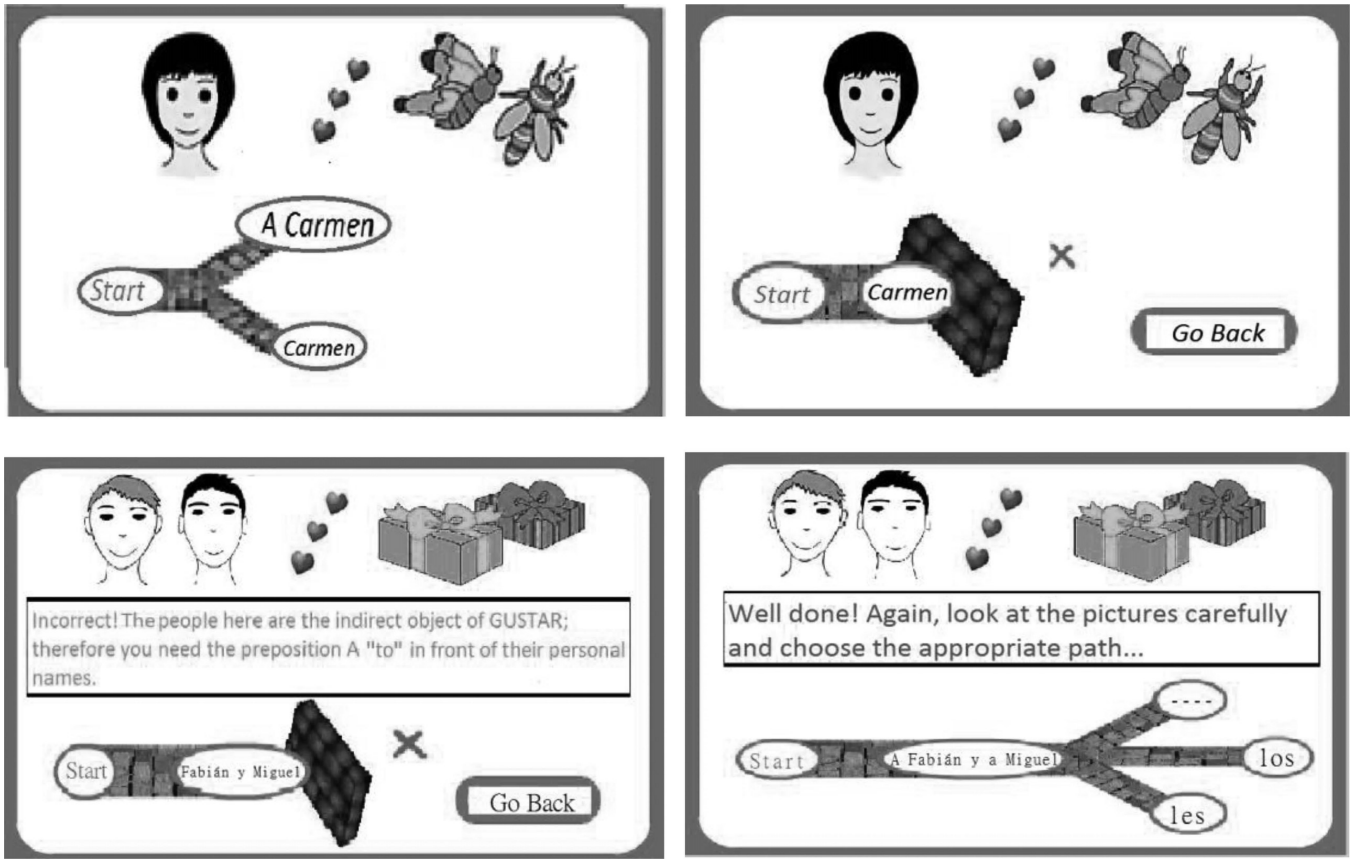

Figure 9.1 Illustration of the task in the implicit and explicit feedback conditions

Table 9.1 Summary of experimental conditions

	No Feedback	*Implicit Feedback*	*Explicit Feedback*
+Practice	PRONLY (Practice only)	PRIFB	PREFB
–Practice (Exposure-to-practice)	EXPONLY (Exposure only)	EXPIFB	EXPEFB

Assessment Tasks

The original study employed a pretest-posttest-delayed design to identify the impact of practice and feedback on L2 learners' *gustar* development. A split-block design was used: Three versions of test were created, and three tests (oral, written, and recognition; see Appendix B for examples) were administered to every participant in a scrambled order to avoid the order and practice effects. The verb *amar* ('to love,' a transitive verb like the English verb) was included as a distractor in these tests. New items with the same level of difficulty were created in order to avoid memory effects. There were 20 items for each of the three tests, including 12 target items and eight distractors.

Different from many conventional recognition tests, the "None of the above"[2] option was included as the fifth one in the multiple-choice questions in order to avoid any potential "learning" by doing the test and to cover all other ill-formed expressions to keep test length short. Furthermore, some test items were in a negative form (such as 'No me gusta para nada la torta') while some were in different degrees of likeness (such as 'Me gusta mucho la torta/No me gusta mucho la torta'), which did not elicit different uses of *gustar* expression so as to avoid monotonous repetition (see Appendix B items 8 and 9). For the oral test, each picture was shown for 15 seconds, and meanwhile the participants recorded their utterance describing the picture. The recordings showed that they took their time easily producing orally. The oral response was later transcribed and scored. For the written and recognition tests, time was not controlled, but the participants completed them with no delay.

Procedure

In the first session of the experiment, participants filled out the Institutional Review Board consent form and the personal information and completed the pretest. For the oral pretest, the participants were exposed to pictures and had to record their oral expression of *gustar* using *Audio Hi-jack*.

In the second session of the experiment, an average of eight days later, participants reported to the language laboratory and were randomly assigned to one of the six experimental groups. All participants were first required to read an explanation of what they should do in the experiment, and then they clicked 'Continue' to proceed. Participants were to use *gustar* to describe pictures, and the computerized game led them through the dative experiencer, indirect object pronoun, verb, and the subject step by step to form the complete expression of *gustar* according to the learning condition to which they were randomly assigned. They were also requested to think aloud as they performed the pathway. Upon completion of the experimental task, they were requested to do the immediate posttest.

In the third session of the experiment, about two and a half weeks later, participants were given once more another set of computerized delayed posttests consisting of oral, written, and recognition to address whether any learning effects were retained. In addition, a debriefing

questionnaire designed to control for any potential outside exposure to the targeted structure was administered as well. The results of this questionnaire revealed that many participants checked the use of the *gustar* structure after the immediate posttest. Thus, the delayed posttest was subsequently discarded from the statistical analyses.

Coding Procedures

To investigate the construct of DoP, the concurrent think-aloud protocols were transcribed and revealed that many of the participants demonstrated some prior knowledge of the verb forms *gustar*; thus, only the DoP counts of preposition *a* and indirect object pronoun (*me, le*, and *les*) were included for the analysis. The protocols were coded by two raters, who first discussed the criteria and then coded independently and counted the instances of verbal reports corresponding to a particular level or depth of processing. To reach 100% agreement on the coding, the raters met to agree upon the coding assignment. For this study, there were three levels of DoP (low, medium, or high) following the coding scheme created by Leow (2015) for grammatical items (p. 228).

A low level of DoP was assigned to participants whose protocols showed no potential for processing target form grammatically. They just read or repeated target items quickly. Protocols that showed comments on a target item in relation to grammatical features, however, were coded as a medium level of DoP. Finally, a high level of DoP was assigned to protocols revealing that participants arrived at a partially or fully accurate or inaccurate underlying grammatical rule.

The following sample protocols illustrated different depths of processing for dative experiencer and indirect object pronoun of the first person (I) in *gustar* expression.

(1) A low level of DoP:

"yo me . . . gusta la torta. . .," " A mí . . . blank . . . like . . . gusta . . . la sandía. . . ," "yo . . . gusta las cerezas." (PRONLY and EXPONLY)

"a mí me gusta la sandia, a mí me gustan las cerezas" (PRIFB, EXPIFB, PREFB, and EXPEFB)

(2) An intermediate level of DoP:

"yo, me, gusta, I don't know what the me means but I picked it, I think that's right, uh, I don't know what a mí means, so I am just gonna pick the yo." (PRONLY)

". . . three pictures, someone likes tennis, that's yo y a mí . . . No . . . a mí . . . okay . . . Then you have to pick between blank and me, but that's probably blank, it's been me, okay." (PRIFB)

(3) A high level of DoP:

"What? What? Okay, I am clicking on yo because it's me, liking tennis, but that is wrong, so that's a friend, so . . . there is no one else, but that's wrong, alright so a friend, me . . . I don't . . . who else could be . . . cause there is only one person in the picture, um, if there is someone speaking, then it's gusta, uhm, it's gonna be yo because I assume we are talking about me, but no, I guess we are not, okay . . . a mí . . . there is no . . . oh, a mí is me, uh, and me is the direct object." (PRIFB)

"yo, because I am talking about me, I love tennis, yo . . . I think a mí is French, so I will definitely stick with yo, ooh, ooh, wrong, 'gustar the English subject is the indirect object, *murmuring* [she was reading the explanation].' Oh, a mí, okay! Two words, I thought that was one word. I don't understand the top choice, just the dash lines or a mí . . .

> me . . . I would probably choose the dash lines because we already have the personal pronoun, guess not, I was wrong [reading the explanation] . . . so we have an indirect object pronoun, okay" (EXPEFB)

These protocols were gathered when participants were processing the *gustar* structures of a + dative and indirect object pronoun + *gusta*. Participants seemed to process the structure with different degrees of cognitive load according to the learning conditions they were assigned in.

Results

To answer the research question, which addressed whether the type of computerized feedback would elicit different levels of processing of the *gustar* structure, a non-parametric test of homogeneity proportions, a 3×3 chi-square (type of feedback × depth of processing) was conducted to calculate the frequency of processing per level vs. per type of feedback. All instances of participants' depth of processing for *gustar* during the treatment was counted into frequencies of low, medium, and high levels (see Table 9.2).

The 3×3 chi-square (type of feedback × DoP level) analysis showed significant results, with χ^2 (4) = 23.934, p = .000 (see Table 9.3); the Contingency Coefficent (CC = .107) presented a moderate effect size (see Table 9.4), and then a post hoc Z test was conducted to locate where the difference was by using the subscript letter that denotes a subset of the categories whose column proportions did not differ significantly from each other at the 0.5 level (see Table 9.5).

Table 9.2 Number of instances of DoP per level and per type of feedback

Feedback	*N*	*DoP*			
		Low	*Medium*	*High*	*Total*
No Feedback	18	202	46	1	249
Implicit	19	624	95	40	759
Explicit	26	895	108	60	1063
Total	63	1721	249	101	2071

Table 9.3 Chi-square tests

	Value	*df*	*Asymp. Sig. (2-sided)*
Pearson Chi-Square	23.934[a]	4	.000
Likelihood Ratio	29.877	4	.000
Linear-by-Linear Association	.028	1	.867
N of Valid Cases	2071		

[a]. 0 cells (.0%) have expected count less than 5. The minimum expected count is 12.14.

Table 9.4 Symmetric measures

		Value	*Approx. Sig.*
Nominal by Nominal	Contingency Coefficient	.107	.000
N of Valid Cases		2071	

Table 9.5 DoP by feedback cross-tabulation

			DoP			*Total*
			Low	*Medium*	*High*	
FB	No Feedback	Count	202_{a}	46_{b}	1_{c}	249
		Expected Count	206.9	29.9	12.1	249.0
		% within FB	81.1%	18.5%	.4%	100.0%
		% within DoP	11.7%	18.5%	1.0%	12.0%
		% of Total	9.8%	2.2%	.0%	12.0%
	Implicit	Count	624_{a}	95_{a}	40_{a}	759
		Expected Count	630.7	91.3	37.0	759.0
		% within FB	82.2%	12.5%	5.3%	100.0%
		% within DoP	36.3%	38.2%	39.6%	36.6%
		% of Total	30.1%	4.6%	1.9%	36.6%
	Explicit	Count	895_{a}	108_{b}	60_{a}	1063
		Expected Count	883.4	127.8	51.8	1063.0
		% within FB	84.2%	10.2%	5.6%	100.0%
		% within DoP	52.0%	43.4%	59.4%	51.3%
		% of Total	43.2%	5.2%	2.9%	51.3%
Total		Count	1721	249	101	2071
		Expected Count	1721.0	249.0	101.0	2071.0
		% within FB	83.1%	12.0%	4.9%	100.0%
		% within DoP	100.0%	100.0%	100.0%	100.0%
		% of Total	83.1%	12.0%	4.9%	100.0%

Note: Each subscript letter denotes a subset of the DoP categories whose column proportions do not differ significantly from each other at the .05 level.

The concurrent data resulted in 63 think-aloud protocols (no feedback: 18, implicit feedback: 19, explicit feedback: 26). A total of 2,071 processing instances were generated under the three experimental conditions. As Table 9.2 shows, overall, both the explicit and implicit feedback elicited higher mean instances of processing than the no feedback (40.88, 39.94 vs. 13.83). The post hoc Z test revealed that the no feedback group elicited significantly different percentages of processing instances within DoP (medium > low > high; 18.5% vs. 11.7% vs. 1.0%.); in contrast, the implicit feedback group elicited similar amount of instances of processing within DoP (38.2% vs. 36.3% vs. 39.6%, high), while the learners in the explicit feedback group produced significantly more instances of low and high DoP than the medium DoP (52.0%, 59.4% vs. 43.4%). From these numbers, it appears that the no feedback group elicited statistically significant variations of processing instances among levels, while both feedback groups elicited statistically less variations between levels of DoP.

Discussion

The results for the role feedback plays in learners' depth of processing appear to indicate that the participants who received feedback, implicit or explicit, spent more cognitive effort on the two of the *gustar* structure than those from the no feedback group. In addition, while the cogni-

tive effort demonstrated in the no feedback group varied significantly with respect to the three different processing levels, this variation was not found in the implicit and explicit feedback groups. In other words, the lack of feedback provision led to substantially less cognitive effort and engagement when compared to those in the feedback groups.

The verbal reports also revealed that participants spent a great amount of time processing cognitively at a low level of processing across all groups: No feedback (81.1%), implicit (82.2%), and explicit feedback 84.2%). These results seem to indicate that most learners potentially processed the metalinguistic information provided to them very swiftly without putting too much of their effort into analyzing them, which potentially reduced the role of feedback to confirm or disconfirm their hypotheses. However, feedback seemed to play a significant role in eliciting a higher level of processing (no feedback vs. implicit and explicit feedback, 0.4% vs. 5.3% and 5.6%, respectively) while its explicitness did not play much of a role in terms of eliciting different levels of processing (5.3% vs. 5.6% of high DoP; 12.5% vs. 10.2 % of medium DoP). Nevertheless, these results should be interpreted with caution since the analysis included participants with some low degree of prior knowledge, which can explain the 18.5% and 1% of processing instances of medium and high level, respectively, for the no feedback group in Table 9.5.

The similar performance between the implicit and the explicit feedback groups in Hsieh (2007) can be explained by an analysis of protocols using Leow's (2015) model of the L2 learning process in ISLA. The results appear to indicate that extra metalinguistic information does not seem to encourage learners to process deeper than a simple right or wrong to confirm what they have just processed in a problem-solving task. The number of verbal instances (5.6%) at high DoP (such as making hypotheses, providing underlying grammatical rule or spending much time processing target item) shows that not many participants in the explicit feedback group benefited from the metalinguistic information offered to them. Indeed, many participants in the explicit feedback condition either glanced over the explanation or just simply skipped it to move along in the puzzle game by simply reading or repeating the target items shown to them, which resulted in a great number of instances of a low depth of processing. The participants in the implicit feedback group, however, without metalinguistic explanation to assist them to figure out why their choices of linguistic items were right or wrong, reported a similar amount of protocols at a high depth of processing and even a few more comments at medium-level processing on the target items given the confirmation of implicit feedback. These concurrent data, then, appear to provide a plausible reason for the non-significant performances of these two feedback groups on the oral, written, and recognition tests in addition to the findings reported in Cerezo (2010), Hsieh, Moreno, and Leow (2016), Sanz and Morgan-Short (2004), and Sauro (2009) (see also Moreno [Chapter 10], this volume, who reported superior DoP for the implicit feedback group). Note that all of these studies included participants with some or substantial amount of prior knowledge of the target items.

Studies (e.g., Bowles, 2008; Rosa & Leow, 2004) that have found significant differences in types of feedback first investigated relatively complex linguistic structures and, second, included participants with very minimal to no prior knowledge of target items. Participants with low prior knowledge might potentially need to process deeper or spend more cognitive effort to encode the target items, thereby heavily relying on explicit information provided to them. Their studies found positive effects for more explicit metalinguistic feedback over implicit feedback. Interestingly, Nagata (1993) and Nagata and Swisher (1995) included participants with substantial prior knowledge, but their participants did not rely any less on the more explicit feedback for this reason. Output production and linguistic complexity may play some role in this case. It was highly potential that their participants experienced an overload in cognitive processing given the output production instead of input practice task and the complex linguistic targets that left them

with fewer processing resources available to reflect on the L2 data. Therefore, the more explicit the feedback, the higher the chances were for participants to benefit from it. More analyses of online think-aloud protocols may provide some insights to explain inconsistent findings.

Pedagogical Implications

The results of the present study suggest that the provision of concurrent computerized feedback to a task that engages learners to actively process L2 input encourages learners to spend more cognitive effort on processing the target L2 data than when they are given no feedback. In addition, the insignificant difference among types of feedback seems to indicate that foreign/second language teachers may not need to provide metalinguistic explanation to a great extent to the learners who have possessed some degree of prior knowledge and to a less-complex linguistic item. The findings in this study do suggest that teachers should design computerized tasks or activities accompanied with concurrent feedback that promote learner processing cognitively during the initial stages of learning. Likewise, teachers need to encourage learners to focus their attention to and process any kind of feedback or error correction provided in the classroom, be it internally driven or externally driven by another student. Given the overall low depth of processing demonstrated by both feedback groups, perhaps the use of prompts, such as questions directing learners' attention to relevant aspects in the input, as employed in Cerezo et al. (2016), may also be implemented in the designing of the computerized task.

Conclusion and Future Directions

In an attempt to shed light on how L2 learners process type of computerized feedback in relation to their performances on an immediate posttest, the present study revisited Hsieh (2007) to code the online think-aloud protocols from a depth of processing perspective. The results suggest that concurrent computerized feedback, irrespective of explicitness, provided in a task that engages learners (with some prior knowledge) to actively process L2, appears to promote significantly greater amount of cognitively high-depth processing. This high depth of processing, in turn, when compared to no feedback, seems to be sufficient to contribute to L2 (*gustar*) development; that is, while computerized practice did not seem to play a role, computerized feedback during practice did, given the large partial eta squared and observed power values obtained on all three types of post-exposure performances and the moderate effect size on the tests for the effects of feedback. Finally, future investigation is warranted before making any definitive statement on the effects of type of computerized feedback. Different types of target form or structures, proficiency levels, L2s, and even other L1s are also fruitful variables to be explored in relation to computerized feedback.

Notes

1 Data from a pilot study revealed that 30 linguistic items (i.e., five items per each set of *gustar* expression) in a treatment task were too long for several participants. In addition, some participants had correlated the direct object pronoun (la/lo/las/los) or indirect object pronouns (le/les) with the object that did the pleasing instead of the person who was pleased.

2 A total of 64.8% of the participants in the present study had chosen this option at least once.

References

Adrada-Rafael, S. (2017). Processing the Spanish imperfect subjunctive: Depth of processing under different instructional conditions. *Applied Psycholinguistics*, *38*, 477–508. doi:10.1017/S0142716416000308

Bowles, M. A. (2008). Task type and reactivity of verbal reports in SLA: A first look at an L2 task other than reading. *Studies in Second Language Acquisition*, *30*(3), 359–387.

Cerezo, L. (2010). *Talking to Avatar: The computer as a tutor and the incidence of practice, feedback, and linguistic form in SLA* (Unpublished doctoral dissertation), Georgetown University, Washington, D. C.

Cerezo, L. (2016). Type and amount of input-based practice in CALI: The revelations of a triangulated research design. *Language Learning & Technology*, *20*(1), 100–123. Retrieved from http://llt.msu.eud/issues/february2016/cerezo.pdf

Cerezo, L., Baralt, M., Suh, B-R., & Leow, R. P. (2014). Does the medium really matter in L2 development? The validity of CALL research designs. *Computer Assisted Language Learning*, *27*(4), 294–310. doi:10.1080/09588221.2013.839569

Cerezo, L., Caras, A., & Leow, R. (2016). The effectiveness of guided induction versus deductive instruction on the development of complex Spanish *gustar* structures: An analysis of learning outcomes and processes. *Studies in Second Language Acquisition*, *38*(2), 265–291. doi:10.1017/S0272263116000139

Chapelle, C. (2003). *English language learning and technology*. Amsterdam, The Netherlands: John Benjamins.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684.

Diez-Bedmar, M. B., & Perez-Paredes, P. (2012). The types and effects of peer native speakers' feedback on CMC. *Language Learning & Technology*, *16*(1), 62–90.

Hsieh, H. C. (2007). *Input-based practice, feedback, awareness and L2 development through a computerized task* (Unpublished doctoral dissertation), Georgetown University, Washington, DC.

Hsieh, H. C., Moreno, N., & Leow, R. P. (2016). Awareness, type of medium, and L2 development: Revisiting Hsieh (2008). In R. P. Leow, L. Cerezo, & M. Baralt (Eds.), *A psycholinguistic approach to technology and language learning* (pp. 131–150). Berlin: Mouton de Gruyter.

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

Jones, C., & Shao, B. (2011). *The net generation and digital natives: Implications for higher education*. York: Higher Education Academy.

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

Leow, R. P. (2007). Input in the L2 classroom: An attentional perspective on receptive practice. In R. DeKeyser (Ed.), *Practice in second language learning: Perspectives from applied linguistics and cognitive psychology* (pp. 21–50). Cambridge: Cambridge University Press.

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. New York, NY: Routledge.

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

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

Loewen, S., & Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning*, *19*(1), 1–14.

Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. M. Gass (Eds.), *Task and language learning: Integrating theory and practice* (pp. 123–167). Clevedon: Multilingual Matters.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–452). Oxford: Oxford University Press.

Montrul, S. (1997). Spanish gustar psych verbs and the unaccusative se construction: The case of dative experiencers in SLA. In A. T. Pérez-Leroux & W. R. Glass (Eds.), *Contemporary perspectives on the acquisition of Spanish, volume I: Developing grammars* (pp. 189–207). Somerville, MA: Cascadilla Press.

Morgan-Short, K., Heil, J., Botero-Moriarty, A., & Ebert, S. (2012). Allocation of attention to second language form and meaning: Issues of think-alouds and depth of processing. *Studies in Second Language Acquisition*, *34*(4), 659–685.

Nagata, N. (1993). Intelligent computer feedback for second language instruction. *The Modern Language Journal*, *77*, 330–339.

Nagata, N., & Swisher, V. (1995). A study of consciousness-raising by computer: The effect of metalinguistic feedback on second language learning. *Foreign Language Annals*, *28*, 337–347.

Prensky, M. (2001, October). Digital natives, digital immigrants. *On the Horizon*, *9*(5), 1–6.

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

Robinson, P., Mackey, A., Gass, S. M., & Schmidt, R. (2012). Attention and awareness in second language acquisition. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 247–267). London: Routledge.

Rosa, E., & Leow, R. P. (2004). Computerized task-based exposure, explicitness, type of feedback, and Spanish L2 development. *The Modern Language Journal*, *88*, 192–217.

Rott, S. (2005). Processing glosses: A qualitative exploration of how form-meaning connections are established and strengthened. *Reading in a Foreign Language*, *17*, 95–124.

Sanz, C. (1999). What form to focus on? Linguistics, language awareness, and the education of L2 teachers. In J. F. Lee, A. Valdman, & S. S. Magnan (Eds.), *Form and meaning: Multiple perspectives* (pp. 3–23). Boston, MA: Heinle & Heinle.

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.

Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning and Technology*, *13*, 96–120.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex.

Whitley, M. S. (1995). *Gustar* and other psych verbs: A problem in transitivity. *Hispania*, *78*, 573–585.

Yilmaz, Y. (2012). Relative effects of explicit and implicit feedback: The role of working memory capacity and language analytic ability. *Applied Linguistics*, *34*(3), 344–368. https://doi.org/10.1093/applin/ams044

Yun, S., Miller, P. C., Baek, Y., Jung, J., & Ko, M. (2008). Improving recall and transfer skills through vocabulary building in web-based second language learning: An examination by item and feedback type. *Educational Technology & Society*, *11*(4), 158–172.

Appendix A

Summary of ANOVAs for the oral, written production, and recognition tests

Source	*SS*	*Df*	*MS*	*F*	*Sig*	*PES*	*OP*[a]
Oral Production							
Practice	103.16	1	103.16	1.39	.243	.02	.21
Feedback	556.61	2	278.31	3.75	.029	.12	.66
Practice x Feedback	77.88	2	38.94	.53	.594	.02	.13
Error	4225.68	57	74.14				
Time	5829.77	1	5829.77	125.41	.000	.69	1.00
Time x Practice	133.64	1	133.6	2.88	.095	.05	.39
Time x Feedback	614.65	2	307.32	6.61	.003	.19	.90
Time x Practice x Feedback	196.76	2	98.38	2.12	.130	.07	.42
Error(Time)	2649.62	57	46.48				
Written Production							
Practice	70.99	1	70.99	.78	.381	.01	.14
Feedback	614.73	2	307.37	3.37	.041	.11	.61
Practice x Feedback	90.14	2	45.07	.50	.612	.02	.13
Error	5194.52	57	91.13				
Time	6865.75	1	6865.75	152.40	.000	.73	1.00
Time x Practice	67.69	1	67.69	1.50	.225	.03	.23
Time x Feedback	884.53	2	442.27	9.82	.000	.26	.98
Time x Practice x Feedback	163.52	2	81.76	1.82	.172	.06	.36
Error(Time)	2567.94	57	45.05				
Recognition							
Practice	28.49	1	28.49	.31	.58	.01	.08
Feedback	678.27	2	339.14	3.63	.033	.11	.65
Practice x Feedback	58.13	2	29.07	.31	.73	.01	.10
Error	5318.82	57	93.31				

Source	*SS*	*Df*	*MS*	*F*	*Sig*	*PES*	*OP*[a]
Time	4750.31	1	4750.31	92.23	.000	.62	1.00
Time x Practice	80.99	1	80.99	1.57	.215	.03	.23
Time x Feedback	713.16	2	356.58	6.92	.002	.20	.91
Time x Practice x Feedback	60.20	2	30.10	.58	.56	.02	.14
Error(Time)	2935.68	57	51.50				

[a] Computed using alpha = .05 PES: Partial Eta Squared OP: Observed Power

Appendix B

Oral Production

Instruction: Look at the drawings provided below and describe orally in a complete sentence in Spanish, *using all the words provided, making the correct conjugations, and add other words if it is necessary in order to describe what each picture depicts*. The sentences containing *yo* ("I") refer to yourself. Please speak loudly and clearly into the microphone.

Written Production

Instruction: Look at the drawings provided below and write in a complete sentence in Spanish, *using all the words provided, making the correct conjugations, and add other words if it is necessary in order to describe what each picture depicts*. The sentences containing *yo* ("I") refer to yourself.

Recognition

Instruction**:** Look at the following drawings and then click on the letter that best describes each picture. Make sure you read through all the options and do your best at selecting the right answer!