Using Processing Instruction to Teach English Reflexives to ESL Learners: A Computer Delivered Study [1]

Rong Liu
English Department, Georgia Gwinnett College, USA

Abstract

The study examines the effects of using computer delivered Processing Instruction (PI) to teach English reflexives. Thirty intermediate ESL learners participated in the pretest-treatment-posttest study. Participants received the input-based PI activities. Gains were assessed by traditional offline tasks such as sentence interpretation and production tasks as well as a psycholinguistic online task, i.e., a self-paced reading task. Results showed that the input-based Processing Instruction was effective in improving participants' ability to both interpret and produce the target forms as measured by traditional offline tasks. More importantly, the results also suggest that PI was effective as measured by psycholinguistic self-paced online reading tasks.

Keywords: Computer delivered, CALL, processing Instruction, reflexives, offline measurement, self-paced online reading tasks, SLA, ESL

Many studies (e.g., Benati, 2021; Naami & Sahragard, 2022; Shintani, Li & Ellis, 2013; VanPatten, 2004; VanPatten & Cadierno, 1993) have found that Processing Instruction (PI), an input-based pedagogical technique, is superior to traditional instruction. PI is one of the teaching methods that take learners' psycholinguistic strategies into consideration for the purpose of instruction (Lee & Benati, 2009; VanPattern, 1996, 2004). For example, leaners have difficulties acquiring the simple past morpheme "-ed" and one reason is because they tend to use lexical information rather than grammatical information to process language input. When learners read "I cooked noodles at home yesterday", learners tend to ignore what the grammatical morpheme "-ed" means and rely on the lexical word "yesterday" to understand the sentence. To help learners acquire the grammatical morpheme "-ed", instruction needs to force learners to process the grammatical information in the morpheme "-ed" in the word "cooked". In a typical PI activity, learners could be asked to read a sentence such as "I cooked noodles at home" and then identify if the action is taking place now or in the past. Without the lexical word "yesterday", learners now have to rely on the morpheme "-ed" to process the input and answer the question. PI is one of the teaching techniques that help learners build the form-meaning connections between grammatical information and meaning. However, it is still inconclusive whether processing instruction can lead to gains in the development of implicit knowledge (i.e., the skills learners can use or apply to process language input on a moment by moment basis) because most previous studies used only offline methods to measure the outcome, which may just capture the gains in explicit knowledge of grammar (i.e., grammatical rules or knowledge that learners can articulate). Ellis (1999) summarized the results of early PI studies and called for further research to clarify whether PI affects implicit knowledge or just raises noticing and understanding

because research has not demonstrated that PI leads to gains in spontaneous use of target structures. Most previous studies used offline measures (such as interpretation tests in which participants listen to sentences and choose the corresponding pictures to represent what they hear, and production tests in which participants complete sentences using given words) for the assessment of the learning effects.

Doughty (2004) also pointed out that offline measurements may just capture metalinguistic or conscious knowledge of the grammar. Currently, the situation has not changed. According to Shintani (2015), out of 42 experiments in 33 published studies involving PI from 1993 to 2013, no studies have incorporated online measures. Psycholinguistic methods are ideal candidates to measure implicit knowledge. The use of explicit knowledge by learners in psycholinguistic tasks such as self-paced reading is minimized (Jiang, 2004, 2007). VanPatten (2002) and Wong and Ito (2018) called for further study using reaction time measures to further test the effects of PI. The current study will employ both psycholinguistic online tasks and traditional offline methods such as interpretation tasks as assessment measures, aiming to capture both implicit knowledge and explicit knowledge gained due to the procedure.

According to VanPatten (2004), PI aims to alter learners' inappropriate processing strategies that lead them to misinterpret or ignore forms in the input (see VanPatten, 2002 for a detailed discussion). It is achieved by structured activities that make form-meaning connections salient to become intake, i.e., the part of input that is processed by learners. According to the Input Processing framework (VanPatten, 1996, 2004), learners derive only limited linguistic information from input and make certain form-meaning connections but not others.

PI has been used to study structures mainly in Spanish (Morgan-Short & Bowden, 2006). For example, in a meta-analysis of PI studies by Shintani, Li and Ellis (2013), they included 16 studies of Spanish, 9 studies of English, 2 studies of French, 8 studies of Japanese. Among the 9 studies of English, there is no PI work targeting English reflexives. It is important to extend the research to different languages and different structures to test the technique. The current study will fill in the gap by examining the effects of PI on teaching English reflexives.

Previous studies also yielded inconclusive findings about the role of feedback type in language teaching. Several studies (Ellis & Sheen, 2006; Long, 2007; Nicholas, Lightbown & Spada, 2001) found that recasts can be beneficial on acquisition, especially when they are more explicit. However, some studies found that explicit feedback did not benefit learners more than other types of feedback (e.g., Lyddon, 2007). In the context of PI, the role of explicit feedback was found to be inconsistent either. Although VanPatten and Borst (2012) found an advantage of explicit information in German case marking instruction, VanPatten and Collopy (2012) did not demonstrate any advantage of the explicit information in PI for Russian case marking. VanPatten et al. (2013) concluded that explicit information may not be essential in PI but may speed up the learning process depending on the nature of the target form. It is clear that the value of explicit feedback in PI needs to be reexamined. This study fills in the gap by examining the role of explicit feedback and implicit feedback in PI.

Another criticism of previous PI studies is that most studies relied on instructors to deliver the intervention (e.g., Sanz, 2003; Shintani, 2015; Shntani, Li & Ellis, 2013). Experiments using traditional classroom instruction involves teachers, which may introduce confounding variables such as different instructors may teach differently on different days. Even the same instructor may teach differently with the same material. The computer program, instead, delivers the same stimuli to all participants in the same way. Additionally, in a classroom setting, "teacher talk" and "student talk" can both serve as input to learners, making it

difficult to control the amount of input each learner obtains in the classroom. With a computer program, each learner gets the same amount of input. In addition to accuracy data, the use of computer technology in the study permits response times to be recorded so that time on task may be analyzed. Studies have shown that, in general, time on task is positively correlated to learning (e.g., Bloom, 1974; Nikolova, 2002). The use of computers allows control of individual and environmental variables such as teacher differences, the timing of stimuli presentation and timing of feedback. Besides the above advantages as an experimental technique, it also has advantages for learners. Training is individualized since learners control the pace of the lesson. Stimuli may be a combination of audio, visual and text, which is potentially more interesting than in-class lessons. The current study will use a computer-delivered PI to exam the effects of PI using both time sensitive online measurements and traditional offline assessments.

The major purpose of the current intervention study is to test the effectiveness of a computer-delivered PI to teach reflexives, i.e., whether the computer automated instruction leads to interlanguage development as measured by both offline tasks and online time-sensitive tasks. Specifically, research questions are:

1. Are computer-delivered PI effective in teaching ESL learners reflexives?
2. Does the measurement type (offline tasks vs. online tasks) matter in PI?
3. Does the feedback type (implicit vs. explicit) make a difference?

Method

This experiment employed the pretest-treatment-posttest design. As indicated by the research questions, the independent variable is the computer delivered PI treatment in reflexives (with explicit feedback vs. with implicit feedback); the dependent variable, i.e., the effects of PI, is the improvement of learning on reflexives (measured by offline tasks and online tasks). During the pretest, subjects did a self-paced reading task to gauge their online processing strategies, completed traditional paper-and-pencil tests which included an interpretation test and a production test to measure their knowledge about reflexives. For the treatment, participants learned about the grammatical constraints on English pronouns and reflexives through a computer program developed by the researcher using a template designed by the supervisor. The program, using *"Director"*, popular software for CALL, recorded the responses and the length of the time that each subject spent on the tasks. Posttest assessment consisted of an interpretation test similar to the pretest, a sentence completion task (as the production test following most PI studies), and a self-paced reading task. Participants' background information was collected through a questionnaire. All participants are bilinguals with L1 Chinese and L2 English.

Target Form

The target form in this study is reflexives such as himself, and herself. One relevant difference between English and Chinese is that English reflexives typically require a local antecedent whereas Chinese reflexives "ziji" can have either a long-distance antecedent or a local one. For example, in English "herself" only refers to "Flora" in "Mary thought that Flora blamed herself". In the Chinese version of this sentence, "herself" can refer to either "Mary" or "Flora". This means that a Chinese ESL learner must learn that long-distance binding is impossible in English. However, this feature is often not taught explicitly in the classroom or in standard textbooks (Thomas, 1993). Moreover, offline studies (Hirakawa, 1990) found that many ESL learners have not acquired the full properties of anaphora, especially the binding principles,

one of which states that an anaphor (reflexives such as "himself" and reciprocals such as "each other") must be bound in its local domain (Binding Principle A in Chomsky, 1981). In "The mother said that the daughter was preparing herself for the party", English speakers know that "herself" must mean or "co-refer with" "daughter", but in Chinese sentences with the same structure, "herself" can also mean "mother". Therefore, it is difficult for Chinese learners of English, who is not sensitive to process the structural information (Binding Principle A) to correctly interpret such sentences. Chinese learners of English often rely on other semantic or pragmatic information to interpret pronouns and reflexives in sentences. For example, in sentences such as "Mary angrily told me that Susan has spilled a lot of paint on herself", the pragmatic meaning favors "Mary" as the antecedent because she was angry. In Chinese, both "Mary" and "Susan" can be "herself". But the pragmatic meaning will favor "Mary" as the antecedent of "herself". In English, pragmatics does not play a role in the interpretation here and only "Susan" can be the antecedent. Furthermore, reflexives are ideal target forms for the study because previous studies have found that traditional intervention did not help much. For instance, White (1995), administering a 4-week long treatment with three 20-minute sessions each week to 40 ESL learners, showed that the traditional explicit instruction did not improve students' learning of the feature. Processing instruction (PI), which focuses on changing learners' processing strategies, may help learners acquire the feature.

Participants

The participants in the present study were 30 paid volunteers who were adult ESL college students in the United States. The experiment excluded two who received perfect scores on the pretest, rendering the *N* size to 28. All participants' first language was Chinese and they had not fully acquired the English structure that was going to be taught as controlled by a pretest. The survey data (see Table 1) showed they could be classified as intermediate ESL learners. 11 participants who reported their TOEFL score had an average score of 589. All participants' average self-ratings of English proficiency in the four skills (speaking, listening, reading and writing) were 2.7, 2.7, 3.3, 2.8, respectively on a scale of 1 to 5 with 5 the best. Participants' age ranged from 19 to 40, with a mean of 28. Participants started to learn English at an average age of 12, ranging from 7 to 15. All participants were well educated, with a mean number of years of formal education of 18 years, ranging from 12 to 23. Based on the education system in mainland China, this means that on average they had learned English at school for at least 10 years. Participants' mean number of months of residence in the U.S. was 27, ranging from 1 to 72. Their mean years of education in English-speaking countries was 1.5, ranging from 0 to 6. The following table summarizes participants' background information.

Table 1

Participants' Background

	Mean	Minimum	Maximum	SD	N
Age	28	19	40	5.2	28
TOEFL scores	589	510	630	39.0	11
Age starting English	11.9	7	15	1.6	28
Years of formal education	17.8	12	23	3.0	28
Months of residence	26.6	1	72	21.3	28

in the U.S.					
Years of education in English-speaking countries	1.5	0	6	2.0	28
Self-rating of English proficiency					
Speaking	2.7	1	5	0.9	28
Listening	2.7	1	5	0.8	28
Reading	3.3	2	5	0.8	28
Writing	2.8	2	4	0.8	28

Instrumentation

The instruments included offline written tests for the pretest and the posttest, a language profile, self-paced reading tasks for the pretest and the posttest, and computer-delivered processing instruction as the treatment.

Offline Written Test: Pretest and Posttest

The current study employed a multiple-choice comprehension test with biased sentences (see Demirci, 2000; Thomas, 1989, 1993; Yuan, 1998 for similar tests). Sample stimuli are:

Subjects read: After the medical tests were completed, the doctor informed Bill about himself.

Subjects answer: Can "himself" refer to the doctor? Yes/no

Can "himself" refer to Bill? Yes/no

Can "himself" refer to somebody else? Yes/no

The interpretation test consisted of 12 sentences (see Appendix B for test stimuli). There were 9 critical sentences testing participants' knowledge of anaphora. The other three were distractors. A split-block design was used such that one version was used for one group as a pretest and the other was used as a posttest.

The current experiment also had a component of production test, a sentence completion task in which participants read a story and, following the story, completed a sentence with a pronoun which is a natural continuation or response to the story. Some stories were adapted from White et al. (1997) by simplifying the stories and changing the task type from interpretation to production. Participants have to understand the story in order to answer the questions correctly. The following are the sample stories:

Sample critical stimuli

Once a week, Mike used to visit an old woman. On Mike's last visit he saw the old woman point a gun at her head and shoot. The old woman died instantly.

Mike knew that the old woman shot_____.

Sample distractor

Killer Harry was free again. Bill was very scared. Bill called a policeman so the policeman could guard him and make sure he was safe from Killer Harry.

Bill asked the policeman to protect _____.

Computer-Delivered Processing Instruction: Treatment

The researcher created the materials according to the guidelines and principles of PI described in VanPatten (2004) with the aim to change learners' processing strategies from

relying on semantic information and gender information to select the antecedent to the correct strategy used by native speakers, i.e., using structure. The current packet presented 41 instances of the target form in both referential and affective activities such as multiple choice tasks, listening and story reading. The drawings that were paired with the sentences were either modeled after other studies (e.g., White et al., 1997) or created by the researcher. The following are some sample items from the training activities (see Appendix C for the complete activities in the training packet).
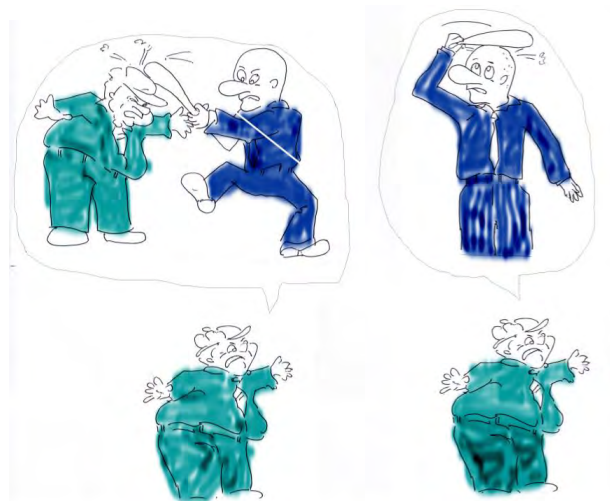
Sample referential activity
Instruction: Listen and then select the picture that goes with the sentence.
Voice stimulus: 1. Mr. Green said angrily that Mr. Blue hit himself.
Visual stimuli: See Figure 1.

Figure 1

Sample referential stimuli

Sample affective activity
Instruction: Select a female classmate or relative or friend of yours (mother, sister, aunt, niece, etc.) and keep her name in mind. Which of the following statements is likely true to her?
1. She says that her father will buy himself a new car. True_____   Not True___

All PI activities were delivered by a computer program called "Director", in which learners followed the lesson by clicking a mouse. In the final training program, learners first logged in using their pseudo names and pseudo-IDs (see Figure 2).
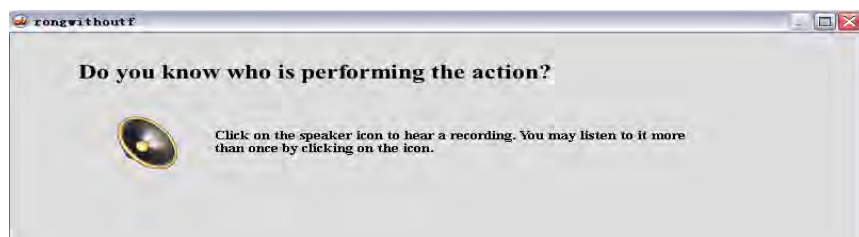
Figure 2

Screen shot: Log in

Then learners clicked the arrow on the lower right screen to continue to the actual training. There were eight activities including matching pictures with recordings, reading sentences and then indicating who is performing the action, true or false affective activities, sentence completion, listening and then indicating the doer of the action, reading stories and then completing sentences, reading passages and answering questions. The following sample screen shots illustrate the program.

First learners saw a screen as shown in Figure 3, and they were given instructions to click on the speaker icon to listen to the recordings. After learners clicked on the speaker, the program played the voice stimulus and then, as soon as it ended, two pictures were displayed on the screen along with instructions asking learners to choose one picture to match the recording. This is shown in Figure 4. When learners clicked on one picture, the program provided immediate feedback as shown in Figure 5. The type of feedback learners received depended on what group they were in, i.e., the explicit feedback group received an explicit explanation and the implicit feedback group was simply told "right" or "wrong" (see Figure 7 for the illustration). Time spent on each event was recorded from the moment when learners clicked on the speaker icon to the moment when they clicked on the arrow on the bottom of the screen. When one part was completed, a screen shot such as shown in Figure 6 appeared, and clicking on the arrow took learners to the next part.

Figure 3

Screen shot: Instructions for activity one

Figure 4

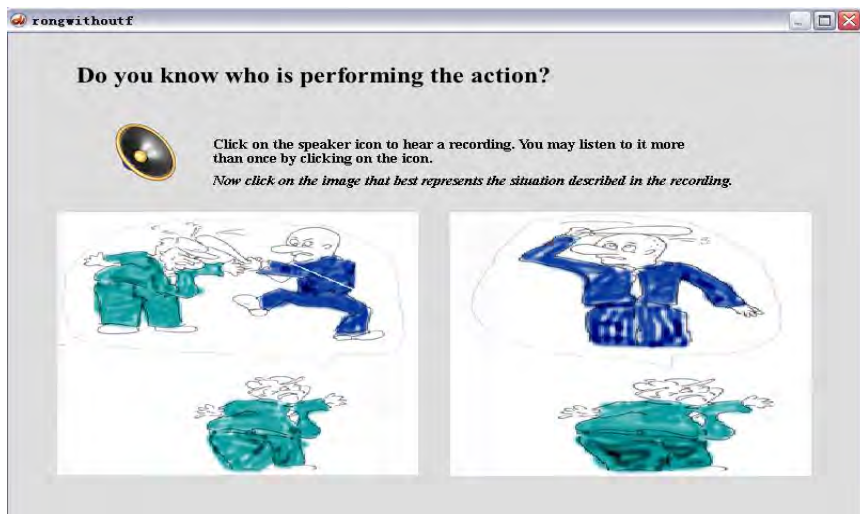*Screen shot: After audio stimulus for activity one*
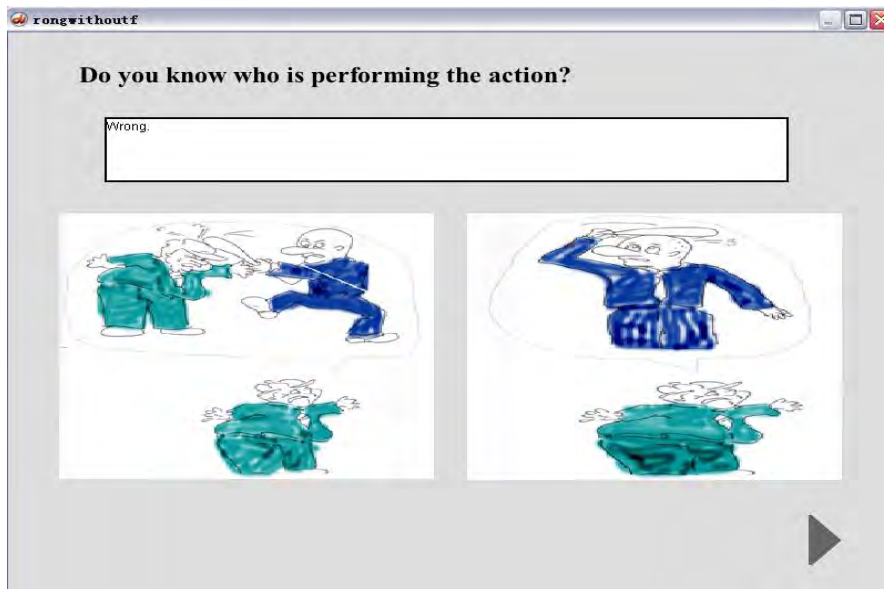

Figure 5

Screen shot: Feedback

Figure 6

Screen shot: End of the section

Figure 7 shows a sample screen shot for the explicit feedback group. This is a referential activity in which learners first listen to a recording and then select a picture that matches the recording. The recording sentence was: "Mr. Green said angrily that Mr. Blue hit himself." This sentence was used to train learners that even if the sentence meaning favors the interpretation that "himself" refers to Mr. Green, the correct response is that "himself" can only refer to the local noun phrase. The explicit feedback for a wrong response would be: "Wrong. Here 'himself' can only refer to Mr. Blue. The sentence meaning may make you choose the other picture, but grammar does not allow that interpretation. Keep in mind that English reflexives such as 'himself' and 'herself' are strictly locally bound, i.e., they refer to the person within the same phrase."

Figure 7

Screen shot: Explicit feedback

Self-Paced Reading Tasks: Pretest and Posttest

Self-paced reading tasks, a task widely used by psycholinguists, were used to assess participants' implicit knowledge of anaphora. Four counter-balanced lists were constructed with each list containing 12 sentences of each type as shown in the following sample stimuli, i.e., Type 1a-4a and Type 1b-4b. Two categories, i.e., pragmatically neutral sentences and pragmatically biased sentences were constructed. For the neutral sentences, the pragmatic meaning of the sentence does not bias toward the choice of the antecedent. For example, in "The son remembered that the father introduced himself at the meeting", the pragmatic meaning of the sentence does not favor either of the two noun phrases "the son" or "the father". Both are plausible. For the biased sentences, the sentence meaning makes the selection of the antecedent structurally incorrect in English. For example, in "The little girl was happy that the mother bought herself a nice toy last week", the sentence meaning favors the interpretation that "herself" means "the little girl" since we know that mothers usually buy toys for their girls and thus the little girl felt happy. Of course, in English grammar, here "herself" corefers to the mother. The sentence construction was modeled after Demirci (2000). Some of the sentences were adapted from previous studies (e.g., Demirci, 2000; Osterhout & Mobley, 1995; Sturt, 2003) and the rest were constructed by the researcher. One novice Chinese ESL learner, who was not a participant in this study, was consulted to see if the sentences rendered the intended effect or confusion, i.e., was asked to interpret the sentences and see if both antecedents were possible or if one was better than the other. One expert in psycholinguistics made necessary changes to make the sentences sound as smooth and natural as possible.

Within each category, there were four types of sentences based on whether the gender cue matches the gender of a potential antecedent and whether the noun phrase (NP) or the potential antecedent is accessible as an antecedent according to Binding Principle A. For example, in the following sample sentences, Type 1a is called a "both match" type because both

the noun phrase "the son" and "the father" are masculine and the gender matches the reflexive "himself". Of course, according to the binding principle, only the second noun phrase "the father" is an accessible antecedent of the reflexive. Other types of sentences are: "only NP2 match" as shown in Type 2a, "only NP1 match" as shown in Type 3a, and "no match" as shown in Type 4a. The noun phrases were half gender specific terms such as father and mother and half common proper names such as Mary and David. Gender neutral terms, such as the chair, were avoided as much as possible. For each category, half sentences used male noun phrases to start the sentences and half used female noun phrases to begin the sentences. No subjects saw all variants of a sentence. Half of the sentences were followed by yes/no comprehension questions. Participants were not probed explicitly about the reflexives in the task so that they were not sensitized to the structure tested explicitly. It was decided that questions should follow some ungrammatical sentences so that participants would not develop a strategy to ignore all ungrammatical sentences. There were 32 filler sentences, serving as distractors (see Appendix A for the test stimuli).

Sample stimuli

Pragmatically neutral sentences:
The son remembered that the father/mother <u>introduced himself/herself at the</u> meeting.
-Was the son introduced at the meeting?
Type 1a: accessible-match/inaccessible match (both match): The son remembered that the father <u>introduced himself at the</u> meeting.
Type 2a: accessible-match/ inaccessible mismatch (only NP2 match): The son remembered that the mother <u>introduced herself at the</u> meeting.
Type 3a: accessible mismatch/inaccessible match (only NP1 match): The son remembered that the mother <u>introduced himself at the</u> meeting.
Type 4a: accessible mismatch/inaccessible mismatch (no match): The son remembered that the father <u>introduced herself at the</u> meeting.
Pragmatically biased sentences:
The little girl was happy that the father/mother <u>bought herself/himself a nice</u> toy last week.
+Was the little girl happy?
Type 1b: accessible-match/inaccessible match (both match): The little girl was happy that the mother <u>bought herself a nice</u> toy last week.
Type 2b: accessible-match/inaccessible mismatch (only NP2 match): The little girl was happy that the father <u>bought himself a nice</u> toy last week.
Type 3b: accessible mismatch/inaccessible match (only NP1 match): The little girl was happy that the father <u>bought herself a nice</u> toy last week.
Type 4b: accessible mismatch/inaccessible mismatch (no match): The little girl was happy that the mother <u>bought himself a nice</u> toy last week.
The same self-paced reading materials were used for the pretest and posttest. To minimize the potential problem of item familiarity between the two tests, posttests were conducted at least two weeks after the pretests.

Procedures
Participants were tested individually, following the following three steps. In Step 1, participants first did the self-paced reading task as a pretest, which took less than 35 minutes. After a 5- to 10-minute break, participants completed the paper-and-pencil test, which assesses

their explicit knowledge of anaphora and it took less than 20 minutes. Participants also filled in a 5-minute language learning background questionnaire during the pretest.

In Step 2, participants were trained through the interactive computer program, which was conducted at least one week after the pretests to minimize the potential sensitizing effect of the pretests. Participants were randomly assigned to two treatment groups, i.e., implicit vs. explicit, based on the order in which they arrived for the experiment. The training took less than 35 minutes for all participants.
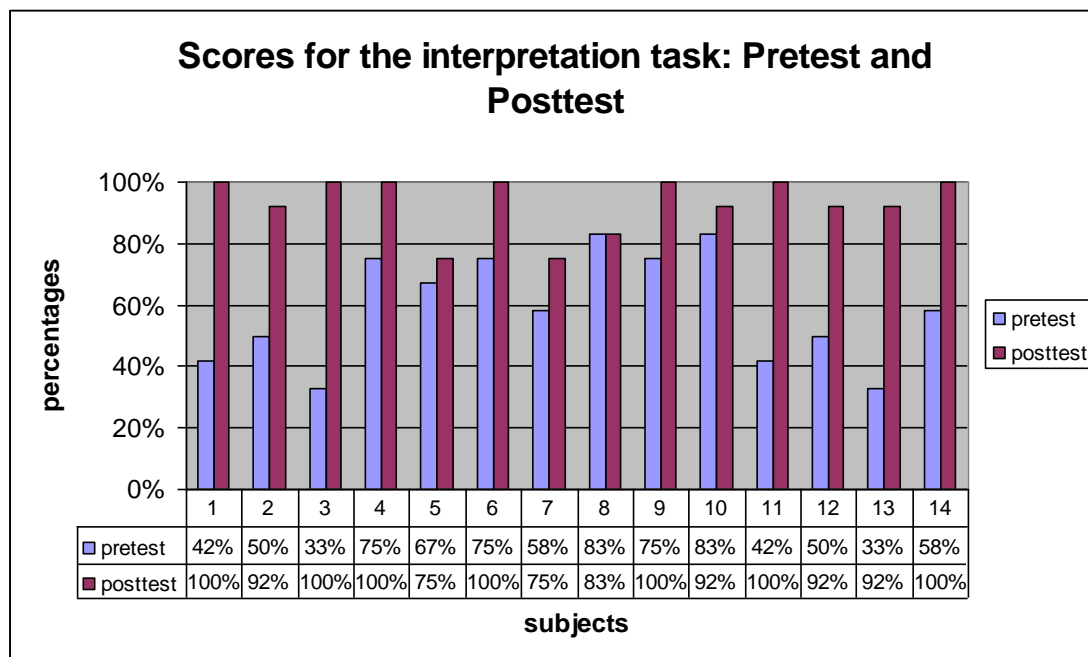
Step 3 was the posttests. Posttests were completed within one week after the treatment. Posttests were not administered immediately after the treatment to minimize the potential item familiarity effect and to determine whether the training had "lasting effects". During the posttest, participants did the self-paced reading task on the computer, which took less than 35 minutes and then took a short break and finished a written test which was a different version of the one they had for the pretest, i.e., if a participant took version A as a pretest, they got version B as a posttest or vice versa. The written test had no time limit but all participants took less than 20 minutes.

Data Analysis and Results

Before examining participants' online reaction time (RT) data, their scores on the written pretest were calculated. There were two parts to the written test, i.e., an interpretation task and a production task. The interpretation task required participants to interpret anaphora by answering three yes/no questions for each item. Those who scored below 95% at the interpretation test were included in the experiment to leave room for participants' improvement (see Sanz & Morgan-short, 2004, for a similar practice). Two participants were excluded for this reason. The same scoring procedure was applied to the posttest. The following figures present the results of the pretest and the posttest for each individual.

Figure 8

Scores of the interpretation task in the pretest and posttest for each individual in the implicit feedback group

Scores for the interpretation task: Pretest and Posttest

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
pretest	42%	50%	33%	75%	67%	75%	58%	83%	75%	83%	42%	50%	33%	58%
posttest	100%	92%	100%	100%	75%	100%	75%	83%	100%	92%	100%	92%	92%	100%

subjects

Note. The mean for pretest was 0.59 and the mean for posttest was 0.93.

Figure 9

Scores of the interpretation task in the pretest and posttest for each individual in the explicit feedback group

Scores for the interpretation task: Pretest and posttest

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
pretest	67%	50%	67%	92%	42%	8%	42%	58%	83%	58%	58%	75%	58%	25%
posttest	92%	100%	100%	100%	100%	92%	100%	58%	92%	83%	100%	92%	100%	50%

subjects
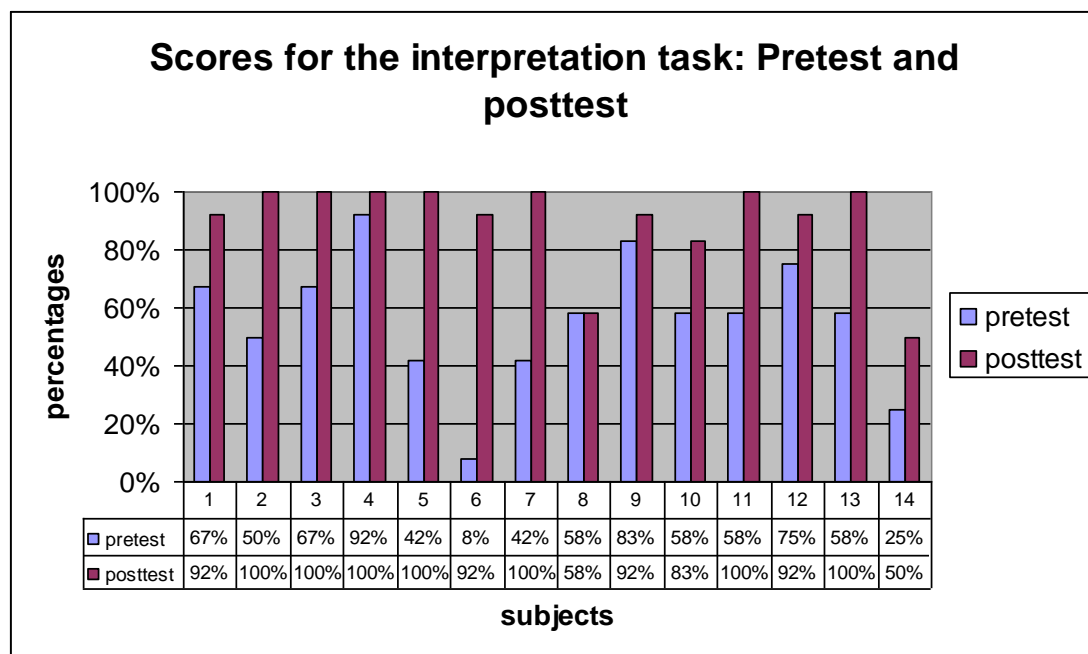
Note. The mean for pretest was 0.56 and the mean for posttest was 0.90.

Figure 10

Scores of the production task in the pretest and posttest for each individual in the implicit feedback group

Scores for the production task: Pretest and Posttest

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
pretest	100%	100%	100%	100%	100%	100%	90%	100%	80%	100%	100%	100%	100%	100%
posttest	100%	90%	90%	90%	80%	100%	100%	90%	90%	100%	100%	100%	100%	100%
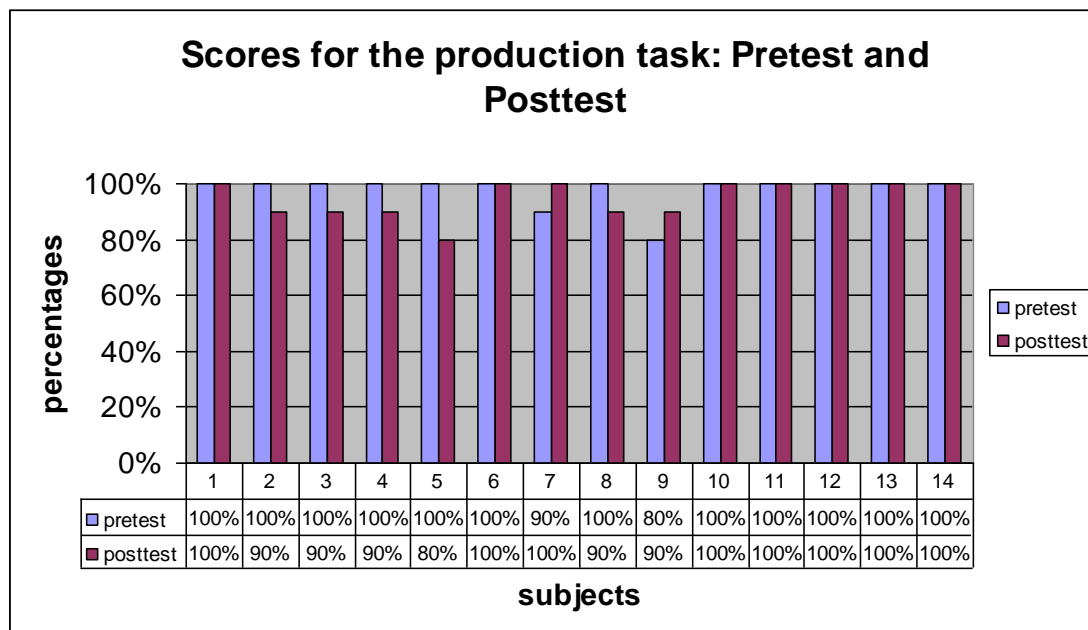
Note. The mean for pretest was 0.98 and the mean for posttest was 0.95.

Figure 11
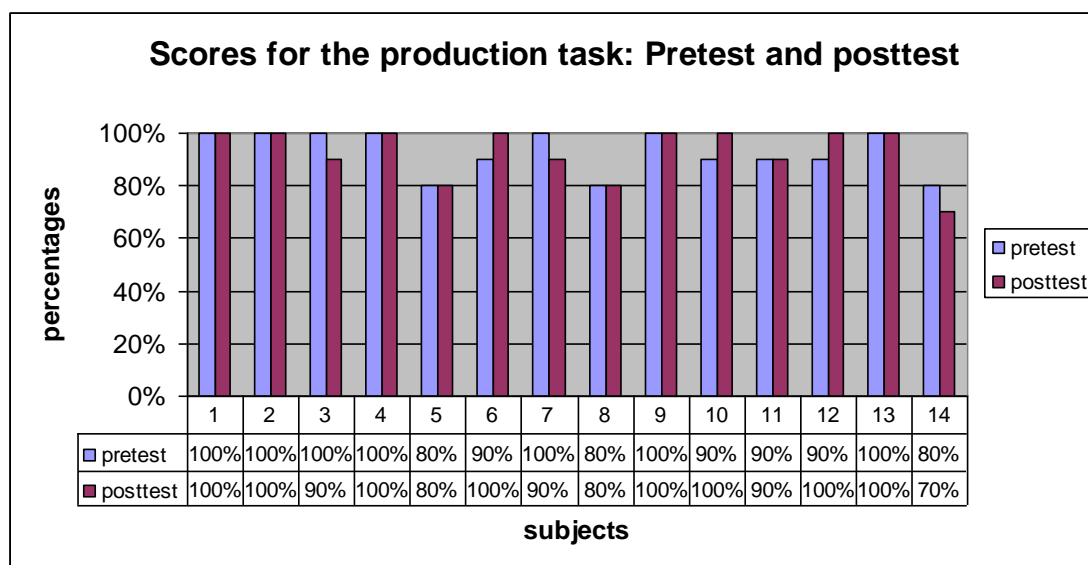
Scores of the production task in the pretest and posttest for each individual in the explicit feedback group

Scores for the production task: Pretest and posttest

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
pretest	100%	100%	100%	100%	80%	90%	100%	80%	100%	90%	90%	90%	100%	80%
posttest	100%	100%	90%	100%	80%	100%	90%	80%	100%	100%	90%	100%	100%	70%

Note. The mean for pretest was 0.93 and the mean for posttest was 0.93.

The data were subjected to repeated measures ANOVA tests in SPSS. Tests of within-subjects contrasts showed that the difference between pretest interpretation and posttest interpretation scores was significant, $F(1, 26)=64.1$, $p<.05$. The difference between pretest production scores and posttest production scores was not significant, $F(1, 26)=1.0$, $p=.32$. Tests of between-subjects effects showed that the difference between the implicit feedback group and the explicit feedback group in improvement in interpretation was not significant, $F(1, 26)=.39$, $p=.54$. The same was true of the production score improvement, $F(1, 26)=2.0$, $p=.17$.

Time on task is a potential variable for the training effect, i.e., the longer participants spend on the task, the better improvement they obtain. The "Director" program was designed to record the time participants spent on each item. For some unknown reasons, data for some participants for some items were not recorded in the output file. Instead of using the total time on the training task to compare the two treatment groups, it was decided to calculate average time spent on each item for each participant. The average time for the implicit group and the explicit group (N=14) was 23 seconds (SD=8) and 21 seconds (SD=6), respectively. Independent Samples T-test showed that the two groups did not differ in the average time they spent on each item, $t=0.6$, $df=26$, $p=0.53$.

The average comprehension error rate in the pretest was 26%, ranging from 15% to 40% (SD=0.07). The mean error rate in the posttest was 21%, ranging from 8% to 40% (SD=0.09). This suggested that, overall, participants understood most of the sentences. Data for the self-paced reading tasks were also analyzed using the SPSS program. Participants' RTs were trimmed before the actual statistical analysis per standard psycholinguistic studies. Those RTs that were two SDs longer or shorter than the same participant's mean, or higher or lower than the high and low cutoffs set at 2,000 and 200 ms, respectively, were excluded. These procedures, along with missing data and display errors, accounted for 10% of the data. Thirty-two mean RTs were computed for each participant, one for each test position (four test positions: positions 1 to 4) in each condition (four conditions: both match, only NP2 match, only NP1 match and no match) for each sentence category (the neutral sentences and the pragmatically biased sentences). The prediction was that any anomaly (in structure or meaning) would produce a slowdown. Therefore, the statistical test was deemed one-tailed. For the item analysis, sixteen means were calculated, one for each of the four test positions in each of the four conditions. The participants' RTs for each test position, each condition, and each sentence category are presented in the following tables. Data for the pretest and the posttest are:

Table 2

Participants' mean RTs (ms) at four test positions for neutral sentences and pragmatically biased sentences involving the structural cue for conditions "both match" and "no match" in the pretest

	Neutral				Biased toward NP1			
Test position	1 introduced	2 self	3 at	4 the	1 bought	2 self	3 a	4 nice
Both match	647	661	536	516	626	624	521	542

No match	656	678	573	540	598	636	553	540
Difference	-9	-17	-39**	-24	28*	-12	-32**	2

*significant at .05 in subject analysis.
**significant at .05 in both subject and item analyses.

Table 3

Participants' mean RTs (ms) at four test positions for neutral sentences and pragmatically biased sentences involving the structural cue for conditions "only NP1 match" and "only NP2 match" in the pretest

	Neutral				Biased toward NP1			
Test position	1 introduced	2 self	3 at	4 the	1 bought	2 self	3 a	4 nice
Only NP1 match	666	663	556	518	618	629	526	520
Only NP2 match	665	657	528	525	605	640	539	524
Difference	1	6	28	-7	13	-11	-13	-4

Table 4

Participants' mean RTs (ms) at four test positions for neutral sentences and pragmatically biased sentences involving the structural cue for conditions "both match" and "no match" in the posttest

	Neutral				Biased toward NP1			
Test position	1 introduced	2 self	3 at	4 the	1 bought	2 self	3 a	4 nice
Both match	559	601	482	456	548	599	458	437
No match	556	614	521	461	545	625	510	495
Difference	3	-12	-39**	-5	3	-26	-52**	-58**

**significant at .05 in both subject and item analyses.

Table 5

Participants' mean RTs (ms) at four test positions for neutral sentences and pragmatically biased sentences involving the structural cue for conditions "only NP1 match" and "only NP2 match" in the posttest

	Neutral				Biased toward NP1			
Test position	1 introduced	2 self	3 at	4 the	1 bought	2 self	3 a	4 nice
Only NP1 match	580	607	500	472	538	595	479	470
Only NP2 match	555	622	465	451	536	601	469	439
Difference	25	-15	35**	21	2	-6	10	31*

*significant at .05 in subject analysis.
**significant at .05 in both subject and item analyses.

Paired-Samples T-Tests were used to compare participants' mean RTs for four conditions at four different positions and for two sentence categories. For each group, sixteen pairs were compared for subject analysis and item analysis, respectively. The details of the data can be found in Appendix D.

As can be seen from the above tables, in the pretest stage, the participants' RTs at the first two positions showed no reliable differences for nonnative speakers (NNSs) for both neutral and pragmatically biased sentences with the exception of the first position for the biased sentences. However, the difference was only significant by subject analysis but not significant by item analysis. This was expected because participants read essentially the same words in the sentences until they reached the word in the second position. In position 2, according to previous studies, RTs were expected to be the same due to the spill-over effect in self-paced reading tasks. The results at the first two positions again confirmed the no-difference predictions.

RTs at the third and fourth positions showed different patterns for different conditions tested. RTs between the "both match" and "no match" sentences for position 3 for both the neutral sentences and the biased sentences were statistically significant both in subject analysis, $t_1(27)=-2.0$, $p<.05$ for the neutral sentences, $t_1(27)=-2.2$, $p<.05$ for the biases sentences, and in item analysis, $t_2(47)=-2.0$, $p<.05$ for the neutral sentences and $t_2(47)=-1.6$, $p<.05$ for the biases sentences. For the fourth position, the difference was not significant either in subject analysis, $t_1(27)=-1.6$, $p=.06$ for the neutral sentences, $t_1(27)=-.13$, $p=.45$ for the biases sentences or in item analysis, $t_2(47)=-1.1$, $p=.14$ for the neutral sentences and $t_2(47)=-.15$, $p=.44$ for the biases sentences.

The comparison between "only NP1" and "only NP2" sentences yielded no significant results for both neutral and biased sentences in both positions. For neutral sentences, in position 3, the 28 ms difference was found nonsignificant for subject analysis, $t_1(27)=1.6$, $p=.06$ and for item analysis, $t_2(47)=1.2$, $p=.13$. The same was true for the fourth position, in subject analysis, $t_1(27)=-.44$, $p=.33$, and in item analysis, $t_2(47)=-.24$, $p=.41$. And for biased sentences, it was found that in position 3, the 13 ms difference was not significant in subject analysis, $t_1(27)=-.78$,

p=.22, and in item analysis, $t_2(47)$=-.65, *p*=.26. In position 4, it was also found that the 11-ms difference was nonsignificant in both subject analysis, $t_1(27)$=-.22, *p*=.42, and in item analysis, $t_2(47)$=-.14, *p*=.45.

The results for the posttest stage showed a different pattern. As expected, RTs at the first two positions for both neutral and pragmatically biased sentences showed no reliable differences. RTs at the third and fourth positions showed different patterns for the planned comparisons. As in the pretest stage, RTs between the "both match" and "no match" sentences for position 3 for both the neutral sentences and the biased sentences were statistically significant both in subject analysis, $t_1(27)$=-2.5, *p*<.05 for the neutral sentences, $t_1(27)$=-3.4, *p*<.05 for the biases sentences, and in item analysis, $t_2(47)$=-1.9, *p*<.05 for the neutral sentences, and $t_2(47)$=-3.0, *p*<.05 for the biases sentences. For the fourth position, the difference for the neutral sentences was not significant either in subject analysis, $t_1(27)$=-.39, *p*=.35, or in item analysis, $t_2(47)$=-.16, *p*=.44. However, the difference for the biased sentences was found significant in both subject analysis, $t_1(27)$=-3.4, *p*<.05, and in item analysis, $t_2(47)$=-3.3, *p*<.05.

Unlike the pretest, the comparison between "only NP1" and "only NP2" sentences in the posttest yielded a different pattern. No reliable differences were found for positions 1 and 2 for both types of sentences. For neutral sentences, in position 3, the 35 ms difference was found significant for both subject analysis, $t_1(27)$=2.4, *p*<.05, and for item analysis, $t_2(47)$=2.0, *p*<.05. In the fourth position, the 21 ms difference approached significance in subject analysis, $t_1(27)$=1.6, *p*=.07, and was nonsignificant in item analysis, $t_2(47)$=1.3, *p*=.10. And for biased sentences, it was found that in position 3, the 10 ms difference was not significant in subject analysis, $t_1(27)$=.55, *p*=.29, and in item analysis, $t_2(47)$=.32, *p*=.38. In position 4, it was found that the 31 ms difference was significant in subject analysis, $t_1(27)$=1.9, *p*<.05 and approached significance in item analysis, $t_2(47)$=1.5, *p*=.07.

To summarize, as expected, it was found that there were no RT differences in positions 1 and 2 for both the pretest and the posttest stages. In the pretest stage, the only significant RT differences were in positions 3 for both neutral and biased sentences in the comparison between the "both match" and "no match" conditions. No other RT differences were found significant. In the posttest stage, besides the significant RT differences in position 3 for both types of sentences in those conditions, the RT difference in position 3 for the comparison between the "only NP1 match" and the "only NP2 match" conditions for the neutral sentences was significant. And the RT difference in position 4 for the comparison between the "only NP1 match" and the "only NP2 match" conditions for the biased sentences was significant in subject analysis and approached significance in item analysis.

Summary of Findings

The study aimed to answer the following research questions: Are computer-delivered processing instruction effective in teaching English reflexives to ESL learners? Does the measurement type (offline tasks vs. online tasks) matter in PI? Does the feedback type (implicit vs. explicit) make a difference?

The study found that in the pretest participants were not using the appropriate strategies when selecting antecedents. They did slow down when there was no match at all in the sentences for the antecedent as shown in the comparison between the "both match" and "no match" conditions, which was true for both neutral sentences and biased sentences. This suggested that participants have acquired some properties of anaphora. For example, this could mean that they

used the gender information to single out one antecedent and when this failed, they encountered reading difficulties and slowed down their reading process, which is typical when readers have problems making sense of what they are reading. However, they have not acquired the full properties of anaphora, i.e., the pattern of the reading time differences between the "only NP1 match" and the "only NP2 match" conditions showed that they were not utilizing the structural information in the sentences. Sentences in the former condition are ungrammatical and sentences in the latter condition are grammatical. Participants should spend more time reading the ungrammatical version per literature in psycholinguistic reading studies (e.g., Liu & Nicol, 2010). This is not the case in this study. Participants showed no RT differences in all positions tested, which suggested that they treated both types of sentences equally. Those sentences in which the only noun phrase that matched the gender of the anaphora was the inaccessible one were treated as acceptable. This could be attributed to the fact that in their L1 Chinese, such sentences are legitimate.

After the computer-delivered processing instruction, results showed that participants improved their performances significantly as measured by the offline written tests. Most of them disallowed long-distance antecedents and allowed only local binding even under the pressure of pragmatic content favoring the alternatives. Although the production tests did not show evidence of improvement, it may be due to the ceiling effect. Participants had already scored high in the pretest, with an average of 96%. Another possibility as pointed by an anonymous reviewer is that maybe the production test was less challenging compared to the interpretation test since they provided more contextual clues, therefore, the participants might not have shown any improvements. For the interpretation test, participants improved significantly. This is a significant finding since previous intervention studies (e.g., White, 1995) failed to show improvement. Most importantly, the online findings showed that ESL learners, to a certain extent, altered their processing strategies after the treatment. In the posttest, the RT differences were significantly different between the "only NP1 match" condition and the "only NP2 match" condition, which is a critical improvement in comparison to the pretest results. This means that participants have changed their language processing behavior from treating the two conditions as the same to treating them differently. Although there were differences between neutral sentences and biased sentences, i.e., results were more robust for the neutral sentences and less so for the biased sentences. For the biased sentences, the effect appeared rather late, i.e., only in position 4, and was significant in subject analysis but only approached significant in item analysis. This suggests that biased sentences were harder to process for participants. This is understandable because participants have to overcome the pressure of the semantic meaning of the sentences which favors the wrong interpretation, and meanwhile they need to resist the temptation to allow long-distance binding. This finding is significant because it indicates that PI is effective not only as measured by traditional offline tests, but also effective as measured by online psycholinguistic methods.

Results also showed no significant difference in improvement between the group who received explicit feedback and the group who received implicit feedback. Both groups improved in the interpretation task and also improved in the online self-paced reading task. This finding was in line with previous studies (e.g., Lyddon, 2007; Sanz & Morgan-short, 2004).

To summarize, the experiment examined whether a computer-delivered PI could be effective in teaching English reflexives to Chinese-English ESL learners. Findings showed that computer-delivered PI did alter L2 learners' processing strategies and also it improved

participants' performances in the offline tasks. Therefore, the automated computer-delivered instruction was successful.

Discussion

The experiment trained learners with a computer-delivered instruction in the binding principle to see if they can acquire the online processing strategies after the specially designed structured activities. Findings showed that learners improved their performances on the target structure significantly as measured by the offline written tests. Most of them disallowed long-distance antecedents and allowed only local binding. Furthermore, results showed that learners developed the appropriate processing strategies to a certain extent after the treatment. It showed that computer-delivered instruction is an effective pedagogical tool that can be used to teach ESL learners reflexives.

The study adds to the growing literature attesting to the effectiveness of PI. It shows that PI was not only effective as measured by traditional offline tests, but also effective as measured by online tasks. It confirmed previous studies' claims that PI was effective (e.g., Marsden, 2006; Salaberry, 1997; Sanz & Morgan-Short, 2004; Soruc, Qin & Kim, 2017; Takimoto, 2006; Toth, 2006; VanPatten & Sanz, 1995; VanPatten, 2002, 2005).

Although many PI studies demonstrated positive results in teaching grammar, one gap that has not still been filled is whether PI can facilitate the learning of binding principles, which many studies have reported difficult to teach in L2 contexts. For example, White (1995) found that teaching the binding principle was not successful. In her study, the two treatment groups did not improve at all after a four-week-long treatment. Although the failure may be due to the fact that they did not directly teach the rule, i.e., long-distance binding was not allowed, it is still puzzling given that learners had the chance to induce the rule because they were exposed to large numbers of exemplary sentences. Another study by White et al. (1996) taught the long-distance binding directly to a group of Japanese learners. This time they were told that in Japanese the antecedent of "zibun" can be outside of the clause, but they were not taught that the antecedent must be a subject. Results showed that out of twelve participants, two French, one English, two Chinese and two Koreans knew the rule after the treatment. In the current experiment, the implicit group only received yes/no feedback; they improved significantly, as shown by both the offline tests and the online tasks. This group is similar to the reading group in White (1995). The explicit group in the experiment is similar to the other treatment groups in White (1995) and White et al. (1996). Therefore, the success of the current experiment is significant and warrants an explanation.

The difference in instruction may explain the different outcomes. The current experiment used PI, a psycholinguistically based pedagogy, to teach the participants, while the other studies used traditional instruction. As discussed in the literature review, many studies have shown that PI is effective for various target structures. Many studies attested that it was more effective than traditional instruction (TI) (e.g., Benati, 2021). The treatment in studies conducted by White (1995) and White et al. (1996) could be classified as traditional instruction given that they provided learners regular oral and written exercises ranging from mechanical to communicative. The tasks were not structured activities as in the current study. They did not push learners away from inappropriate processing strategies. Learners need to establish form-meaning connections when acquiring a target structure. PI helps learners form new connections and avoid old ones from their L1 that are not appropriate for the new language. One key difference between PI and

TI is that PI aims to help learners get intake from input using design activities that make form-meaning connections between non-salient forms and function salient forms. PI activities also force learners to alter their inappropriate processing strategies for the target language. It emphasizes the important role of input in language acquisition while traditional instruction emphasizes the role of output-based exercises. TI activities do not force learners to change their processing strategies although they do sometimes inform learners about the difference between L1 and L2.

The current study showed that language learners' development in automatic competence and processing routines is selective (Liu & Nicol, 2010) and they are not equivalent to their linguistic competence. Using only offline measurement cannot capture the whole picture of learners' language proficiency. Developing appropriate processing routines or automatic competence is important for learners in order to use or comprehend a second language as efficiently as L1 speakers do.

The findings of the study have some pedagogical implications. Language instruction is more beneficial if instructors focus on the processes that lead learners to turn input to intake, and push learners away from the wrong processing strategies. When instructors keep learners' processing strategies in mind and design activities that are structured to help learners avoid the wrong strategies, language teaching is successful. In the experiment, learning activities were designed according to PI and those structured activities (see VanPatten, 1996 for details) forced learners to focus on processing the input. This change of focus helped learners make new form-meaning connections and made a difference in learners' interlanguage development. According to VanPatten (2004), learners tend to use content words and other lexical items, rather than grammatical forms, to get meaning. Activities should be designed to encourage learners to process grammatical forms in order to correctly understand sentences. This will help learners acquire the correct processing strategies for the target language that can be deployed for automatic processing in spontaneous speech. For example, when teaching reflexives, it is necessary to create sentences that push learners to use the grammatical forms or structures, not lexical items to process the sentences to get correct answers. One such example question "The hungry boy Jack was happy that John brought himself some cakes. Whom did John bring cakes for?" illustrates the strategy employed by PI. Learners usually rely on content words such as *hungry* and *happy* to conclude that *himself* refers to *Jack*. This strategy used by L2 learners is referred as the Primacy of Meaning Principle (see VanPatten, 1996). However, the strategy turns out to be wrong. To understand the sentence correctly, learners must only rely on the grammatical structure, i.e., local binding only in reflexives and ignore the signal from lexical items, even when the meaning of the words favors the other entity *Jack*.

Conclusion

This study used computer delivered PI to teach English reflexives with both offline and online assessments. The results are encouraging for the field of Second Language Acquisition (SLA). Findings showed that computerized instruction, which eliminates many confounding factors in research, can be effective in teaching ESL grammar such as reflexives. The study also showed that the moving-window self-paced reading technique is a valid tool for SLA, especially to investigate the status of knowledge representation, whether explicit or implicit. Results revealed PI is an effective pedagogical technique that can alter learners' inappropriate processing strategies, and lead to gains in interlanguage developing system, not just explicit knowledge.

Although the study showed that PI is effective, it only tested one target structure, i.e., reflexives, in English. Future studies will need to test more structures and more groups of learners with different L1s at different proficiency levels to establish the extent to which the findings can be generalized (see Shintani, 2015 for a meta-analysis of PI studies). Another limitation of the study is the small sample size which might make it difficult to generalize the findings. Although the number of participants was 30, due to the between-subject variable, feedback type, the number of participants in each condition was only 15. Future studies should have more participants. The more use of psycholinguistic methods such as eye-tracking (e.g., Benati, 2021; Laval & Lowe, 2020) to examine the effects of PI intervention on real-time language processing will also shed light on SLA.

References

Benati, A. (2005). The effects of processing instruction, traditional instruction, and meaningful based output instruction on the acquisition of English past simple tense. *Language Teaching Research, 9,* 67-93. https://doi.org/10.1191%2F1362168805lr154oa

Benati, A. (2021). Input Processing and Processing Instruction: The Acquisition of Italian and Modern Standard Arabic. Benjamins Publishing. https://doi.org/10.1075/bpa.11

Bloom, B. S. (1974). Time and learning. *American Psychology, 29*, 682-688. https://psycnet.apa.org/doi/10.1037/h0037632

Chomsky, N. (1981). Lectures on Government and Binding. Foris.

Demirci, M. (2000). The role of pragmatics in reflexive interpretation by Turkish learners of English. *Second Language Research, 16*, 325-353. https://doi.org/10.1177%2F026765830001600402

Doughty, C. J. (2004). Commentary: When PI is focus on form it is very very good, but when it is focus on forms…. In B.VanPatten (Ed.), *Processing instruction: Theory, research, and commentary (*pp. 257-270). Erlbaum.

Ellis, R. (1999). Input based approaches to teaching grammar: A review of classroom-oriented research. *Annual Review of Applied Linguistics, 19*, 64-80. https://doi.org/10.1017/S0267190599190044

Ellis, R. & Sheen, Y. (2006). Re-examining the role of recasts in L2 acquisition. *Studies in Second Language Acquisition, 28*, 575-600. https://doi.org/10.1017/S027226310606027X

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603-634. https://doi.org/10.1017/S0142716404001298

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning, 57*, 1-33. https://doi.org/10.1111/j.1467-9922.2007.00397.x

Juffs, A. (2001). Psycholinguistically oriented second language research. *Annual Review of Applied Linguistics*, *21*, 207–220. https://doi.org/10.1017/S0267190501000125

Laval, C., & Lowe, H. (2020). The use of eye-tracking in experimental approaches in second language acquisition research: the primary effects of Processing Instruction in the acquisition of the French imperfect. *Journal of French Language Studies, 30*(2), 189-210. https://doi.org/10.1017/S0959269520000046

Lee, J. F., & Benati, A. G. (Eds.). (2009). Research and perspectives on processing instruction. Mouton de Gruyter. https://doi.org/10.1515/9783110215335

Lee, J.F., & VanPatten, B. (1995). Making communicative language teaching happen (1st ed.). McGraw-Hill.

Liu, R., & Nicol, J. (2010). Online processing of anaphora by advanced English learners. In Prior, M. T., Watanabe, Y., & Lee, S.-K. (Eds.), *Selected proceedings of the 2008 Second Language Research Forum* (pp. 150–165). Cascadilla Proceedings Project.

Long, M. (2007). Recasts in SLA: The story so far. In M. Long (Ed.), *Problems in SLA* (pp.78-118). Erlbaum.

Lyddon, P. (2007). *The efficacy of corrective feedback and target form enhancement in promoting acquisition of the a/au/en/aux distinction in L2 French*. Doctoral dissertation, The University of Arizona.

Marsden, E. (2006). Exploring input processing in the classroom: An experimental comparison of processing instruction and enriched input. *Language Learning, 56*, 507-566. https://doi.org/10.1111/j.1467-9922.2006.00375.x

Naami, A., & Sahragard, R. (2022). The effect of field independence on processing instruction on the acquisition of English passive by Iranian EFL learners. *Instructed Second Language Acquisition, 6*(1), 109–139. https://doi.org/10.1558/isla.20078

Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning, 51*, 719-758. https://doi.org/10.1111/0023-8333.00172

Nikolova, O. R. (2002). Effects of students' participation in authoring of multimedia materials on student acquisition of vocabulary. *Language, Learning & Technology, 6,* 100-118.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34,* 739-773. https://doi.org/10.1006/jmla.1995.1033

Salaberry, M. (1997). The role of input and output practice in second language acquisition. *Canadian Modern Language Review, 53*, 422-451. https://doi.org/10.3138/cmlr.53.2.422

Sanz, C. (2003). Computer delivered implicit vs. explicit feedback in processing instruction. In VanPatten, B. (Ed.) Processing instruction (pp.241-256). Erlbaum.

Sanz, C., & Morgan-Short, K. (2004). Positive evidence versus explicit rule presentation and explicit negative feedback: A computer-assisted study. *Language Learning, 54,* 35-78. https://doi.org/10.1111/j.1467-9922.2004.00248.x

Shintani, N. (2015). The effectiveness of Processing Instruction and production-based instruction on L2 grammar acquisition: A meta-analysis. *Applied Linguistics, 36,* 306-325. https://doi.org/10.1093/applin/amu067

Shintani, N., Li, S., & Ellis, R. (2013), Comprehension-based versus production-bsed grammar instruction: A meta-analysis of comparative studies. *Language Learning, 63,* 296-329. https://doi.org/10.1111/lang.12001

Slobin, D., & Bever, T. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition, 12,* 229–265. https://doi.org/10.1016/0010-0277(82)90033-6

Soruc, A., Qin, J., & Kim, Y. (2017). Comparing the effectiveness of Processing Instruction and Production-based Instruction on L2 grammar learning: The role of explicit information. *TESL Canada Journal, 34,* 49-70. https://doi.org/10.18806/tesl.v34i2.1266

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory & Language, 48,* 542-562. https://doi.org/10.1016/S0749-596X(02)00536-3

Swain, M. (1995). Three functions of output in second language learning. In G. Cook, & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125-144). Oxford University Press.

Takimoto, M. (2006). The effects of explicit feedback on the development of pragmatic proficiency. *Language Teaching Research, 10*, 393-417. https://doi.org/10.1191/1362168806lr198oa

Thomas, M. (1989). The interpretation of English reflexive pronouns by nonnative speakers. *Studies in Second Language Acquisition, 11*, 281-303. https://doi.org/10.1017/S0272263100008147

Thomas, M. (1993). Knowledge of reflexives in a second language. Amsterdam: Benjamins. https://doi.org/10.1075/lald.6

Toth, P. D. (2006). Processing instruction and a role for output in second language acquisition. *Language Learning, 56*, 319-385. https://doi.org/10.1111/j.0023-8333.2006.00349.x

VanPatten, B. (1984). Learners' comprehension of clitic pronouns: More evidence for a word order strategy. *Hispanic Linguistics, 1*, 57-67.

VanPatten, B. (1996). *Input processing and grammar instruction*. Ablex.

VanPatten, B. (2002). Processing instruction: An update. *Language Learning, 52*, 755-803. https://doi.org/10.1111/1467-9922.00203

VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5-32). Erlbaum. https://doi.org/10.4324/9781410610195

VanPatten, B. (2005). Processing instruction. In C. Sanz (Ed.), *Mind and context in adult second language acquisition* (pp. 267-281). Georgetown University Press.

VanPatten, B., & Borst, S. (2012). The roles of explicit information and grammatical sensitivity in Processing Instruction: Nominative-accusative case marking and word order in German L2. *Foreign Language Annals, 45,* 92-109. https://doi.org/10.1111/j.1944-9720.2012.01169.x

VanPatten, B. & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15*, 225-243. https://doi.org/10.1017/S0272263100011979

VanPatten, B., & Collopy, E. (2012). Explicit information and processing instruction with nominative and accusative case in Russian as a second language: Just how important is explanation? *Slavic and East European Journal, 56,* 256-276.

VanPatten, B., Collopy, E., Price, J. E., Borst, S., & Qualin, A. (2013). Explicit information, grammatical sensitivity, and the First-Noun Principle: A cross-linguistic study in Processing Instruction. *Modern Language Journal, 97,* 506-527. https://doi.org/10.1111/j.1540-4781.2013.12007.x

VanPatten, B., & Sanz, C. (1995). From input to output: Processing instruction and communicative tasks. In D. Eckman, D. Highland, P. Lee, J. Mileham and R. Weber (Eds.), *SLA theory and pedagogy*. Lawrence Erlbaum.

White, L. (1995). Input, triggers and second language acquisition: Can binding be taught? In F. Eckman, D. Highland, P. Lee, J. Mileman & R. Rutkowski Weber (Eds.), *Second language acquisition theory and pedagogy* (pp. 63-78). Lawrence Erlbaum.

White, L., Bruhn-Garavito, J., Kawasaki, T., Pater, J. & Prevost, P. (1997). The researcher gave the subject a test about himself: Problems of ambiguity and preference in the investigation of reflexive binding. *Language Learning, 47,* 145-172. https://doi.org/10.1111/0023-8333.41997004

White, L., Hirakawa, M., & Kawasaki, T. (1996). Effects of instruction on second language acquisition of the Japanese long-distance reflexive zibun. *Canadian Journal of Linguistics, 41,* 235-254. https://doi.org/10.1017/S0008413100016418

Wong, W. & Ito, K. (2018). The effects of processing instruction and traditional instruction on L2 online processing of the causative construction in French: an eye tracking study. *Studies in Second Language Acquisition, 40* (2), 241-268. https://doi.org/10.1017/S0272263117000274

Yuan, B. P. (1998). Interpretation of binding and orientation of the Chinese reflexive ziji by English and Japanese speakers. *Second Language Research, 14*, 324-340. https://doi.org/10.1191/026765898670904111

Appendix A: Online Test Stimuli

1.   The son remembered that the father/mother introduced himself/herself at the meeting.
-Was the son introduced at the meeting?
2.   The man said that the mother/grandson bought herself/himself a small gift from the shop.
3.   John thought that Bill/Mary owed himself/herself another chance to solve the problem.
4.   The cleaning lady knew that the old woman/man shot herself/himself in the hotel.
-Did the old woman/man shoot the cleaning lady in the hotel?
5.   The nurse said that the soldier/policeman felt sorry for himself/herself because of the serious wound.
+/-Did the soldier/policeman feel happy?
6.   The father thought that the daughter/son hated himself/herself because of the incident.
7.   Rose said that Mike/Jill wrote about herself/himself in the letter.
8.   Mary thought that Tom/ Mike blamed herself/himself for being late for the train.
-Was the train late?
9.   Mike believes that Helen/Tom trusts himself/herself to be able to get around New York.
10.   Susan heard that Mary/John had bought herself/himself a new 10-speed bicycle.
+/-Did Mary/John buy a new bicycle?
11.   David could see that Bill/Sue was looking at himself/herself in the mirror.
12.   Sam thinks that Jane/Tom dislikes herself/himself for being so impatient.
-Does Sam dislike Jane/Tom for being so impatient?
13.   Alice knew that John/Mary understood herself/himself pretty well.
14.   Bob said that Paul/Helen hit herself/himself with a long stick.
15.   The waitress said the man/lady criticized himself/herself for being rude at the restaurant.
+/-Was the man/lady rude at the restaurant?
16.   The woman remembered that the daughter/son brought herself/himself a beautiful bouquet of flowers.
+/-Did the daughter/son buy some beautiful flowers?
17.   The little girl was happy that the father/mother bought herself/himself a nice toy last week.
18.   Mary angrily told me that the boy/girl had spilled a lot of paint on herself/himself the other day.
+Did the boy/girl spill some paint the other day?
19.   The little boy was angry that the father/mother hit himself/herself last week at home.

20.   The famous actress heard that the popular host/hostess talked about himself/herself on TV last night.
-Did the famous actress watch the show on TV last night?
21.   The lady said unhappily that the brother/sister often does not trust herself/himself because of the nasty gossip.
22.   The cashier was worried that Mrs. Smith/Mr. Smith wrote herself/himself a bad check.
23.   The student was angry that Rose/Jack gave himself/herself so much homework.
+/-Did Rose/Jack have so much homework?
24.   The boy was excited that the mother/father bought himself/herself a birthday gift.
25.   George felt embarrassed that the woman/man found himself/herself staring at an attractive nurse.
-Was George embarrassed about staring at an attractive nurse?
26.   The talkative uncle was sad that the father/mother kept himself/herself from talking to friends.
27.   The nun was mad that the priest/priestess kept looking at himself/herself after the sermon.
-Was the nun happy with the priest/priestess?
28.   The husband was embarrassed that the lady/guy talked himself/herself into believing the bizarre story.
29.   The hungry boy was very happy that the grandfather/grandmother brought himself/herself a chocolate cake.
30.   The little girl with a bad cold was told that the man/woman would take care of himself/herself during the journey.
+Would the man/woman look after the little girl during the journey?
31.   The grandson was disappointed that the father/mother didn't buy himself/herself a new toy car from the shop.
+Did the grandson feel disappointed about the new toy car?
32.   The new chairwoman was frustrated that the new actor/actress didn't trust himself/herself to finish the big project.
-Did the new actor/actress distrust the new chairwoman?

Appendix B: Offline Test Stimuli

I. Answer each of the questions below (Version A; Subject ID_____)
1. Rose said that Mary voted for herself.
Can "herself" refer to Rose? Yes   No
Can "herself" refer to Mary? Yes   No
Can "herself" refer to somebody else? Yes   No
2. The little boy was angry that the father always hit himself.
Can "himself" refer to the little boy? Yes   No
Can "himself" refer to the father? Yes   No
Can "himself" refer to somebody else? Yes   No
3. Mary heard that the famous actress talked about herself on TV.
Can "herself" refer to Mary? Yes   No
Can "herself" refer to the famous actress? Yes   No
Can "herself" refer to somebody else? Yes   No
4. The man next to the little boy bought himself a new toy from the store.

Can "himself" refer to the man? Yes   No
Can "himself" refer to the little boy? Yes   No
Can "himself" refer to somebody else? Yes   No
5. Jane said that the little girl was dressing herself upstairs.
Can "herself" refer to Jane? Yes   No
Can "herself" refer to the little girl? Yes   No
Can "herself" refer to somebody else? Yes   No
6. Jordon remembered that the popular singer wrote about himself in the autography.
Can "himself" refer to Jordon? Yes   No
Can "himself" refer to the popular singer? Yes   No
Can "himself" refer to somebody else? Yes   No
7. The hungry girl was happy that the mother bought herself an ice cream.
Can "herself" refer to the hungry girl? Yes   No
Can "herself" refer to the mother? Yes   No
Can "herself" refer to somebody else? Yes   No
8. The lady next to Kate introduced herself at the meeting.
Can "herself" refer to the lady? Yes   No
Can "herself" refer to Kate? Yes   No
Can "herself" refer to somebody else? Yes   No
9. Jack says that Peter talked about himself.
Can "himself" refer to Jack? Yes   No
Can "himself" refer to Peter? Yes   No
Can "himself" refer to somebody else? Yes   No
10. The customer was upset that the waiter spilled some soup on himself.
Can "himself" refer to the customer? Yes   No
Can "himself" refer to the waiter? Yes   No
Can "himself" refer to somebody else? Yes   No
11. The patient beside the doctor can look after himself next week.
Can "himself" refer to the patient? Yes   No
Can "himself" refer to the doctor? Yes   No
Can "himself" refer to somebody else? Yes   No
12. The lady said that the old woman blamed herself for being late for the train.
Can "herself" refer to the lady? Yes   No
Can "herself" refer to the old woman? Yes   No
Can "herself" refer to somebody else? Yes   No

II. Read the story and then complete the last sentence with a pronoun or reflexive.
1. Once a week, Mike used to visit an old woman. On Mike's last visit he saw the old woman point a gun at her head and shoot. The old woman died instantly.
Mike knew that the old woman shot_____.
A. him   B. himself   C. her   D. herself
2. Johnny and a little boy were playing with matches. Johnny lit a match and then dropped it on the little boy's leg. The little boy went screaming to his father and told him what had happened.
The little boy said that Johnny had burned_____.
A. him   B. himself   C. her   D. herself

3. Bill was going to a party. A very famous male actor was going to attend the party. Bill was too shy to speak to the actor, so he hoped the actor would speak to him instead.
Bill hoped that the famous actor would introduce_____.
A. him   B. himself   C. her   D. herself
4. Killer Harry was free again. Bill was very scared. Bill called a policeman so the policeman could guard him and make sure he was safe from Killer Harry.
Bill asked the policeman to protect _____.
A. him   B. himself   C. her   D. herself

I. Answer each of the questions below (Version B; Subject ID_____)
1. Bill said that Peter criticized himself for being rude at the restaurant.
Can "himself" refer to Bill? Yes   No
Can "himself" refer to Peter? Yes   No
Can "himself" refer to somebody else? Yes   No
2. The patient was worried that the nurse would spill some medicine on herself.
Can "herself" refer to the patient? Yes   No
Can "herself" refer to the nurse? Yes   No
Can "herself" refer to somebody else? Yes   No
3. Mary believed that the famous actress talked about herself on TV.
Can "herself" refer to Mary? Yes   No
Can "herself" refer to the famous actress? Yes   No
Can "herself" refer to somebody else? Yes   No
4. The son remembered that the father introduced himself at the party.
Can "himself" refer to the son? Yes   No
Can "himself" refer to the father? Yes   No
Can "himself" refer to somebody else? Yes   No
5. The nurse next to Helen can wake up herself the next morning.
Can "herself" refer to the nurse? Yes   No
Can "herself" refer to Helen? Yes   No
Can "herself" refer to somebody else? Yes   No
6. Bill said that the actor was preparing himself upstairs for the performance.
Can "himself" refer to Bill? Yes   No
Can "himself" refer to the actor? Yes   No
Can "himself" refer to somebody else? Yes   No
7. The mother said that the daughter would buy herself a new toy.
Can "herself" refer to the mother? Yes   No
Can "herself" refer to the daughter? Yes   No
Can "herself" refer to somebody else? Yes   No
8. The tourist was angry that the tour guide laughed at himself.
Can "himself" refer to the tourist? Yes   No
Can "himself" refer to the tour guide? Yes   No
Can "himself" refer to somebody else? Yes   No
9. The thirsty boy behind the father bought himself a bottle of water.
Can "himself" refer to the thirsty boy? Yes   No
Can "himself" refer to the father? Yes   No
Can "himself" refer to somebody else? Yes   No
10. Mike was upset that the son made fun of himself in the classroom.

Can "himself" refer to Mike? Yes   No
Can "himself" refer to the son? Yes   No
Can "himself" refer to somebody else? Yes   No
11. The lady beside Mary made herself some new clothes.
Can "herself" refer to the lady? Yes   No
Can "herself" refer to Mary? Yes   No
Can "herself" refer to somebody else? Yes   No
12. The mother said that Susan prepared herself a meal.
Can "herself" refer to the mother? Yes   No
Can "herself" refer to Susan? Yes   No
Can "herself" refer to somebody else? Yes   No

II. Read the story and then complete the last sentence with a pronoun or reflexive.
1. After three years in the war, the soldier finally went crazy and jumped out of a window. He died instantly. The doctor had to tell the soldier's family the sad news.
The doctor said that the soldier killed_____.
A. him   B. himself   C. her   D. herself
2. Susan and her friend were sewing. They were careless and left some pins on the floor. Susan was not wearing shoes and she stepped on a pin. Susan started to shout and cry. Susan's friend could see the blood on Susan's foot.
The friend realized that Susan hurt_____.
A. him   B. himself   C. her   D. herself
3. Mary used to have a problem. Every time she met someone she got nervous and forgot her own name. Mary was going to a party at a friend's house. Mary hoped her friend would help Mary meet people by telling them Mary's name.
Mary hoped that her friend would introduce_____.
A. him   B. himself   C. her   D. herself
4. Johnny and his father were going for a walk. It started to rain. Johnny had a cold, so his father gave him a jacket and told him to put it over his head. Johnny felt warm under the jacket.
His father told Johnny to cover_____ with a jacket.
A. him   B. himself   C. her   D. herself

Appendix C: Computer-delivered Processing Instruction

A. Select the picture that goes with the sentence.
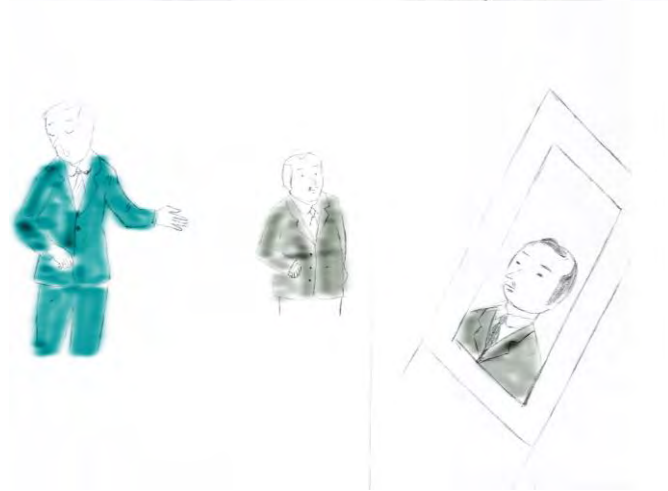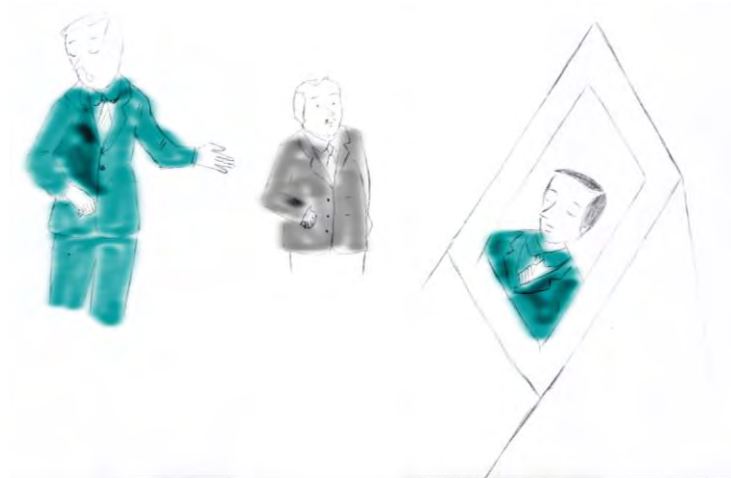   1. Mr. Green explained that Mr. Blue cut himself.
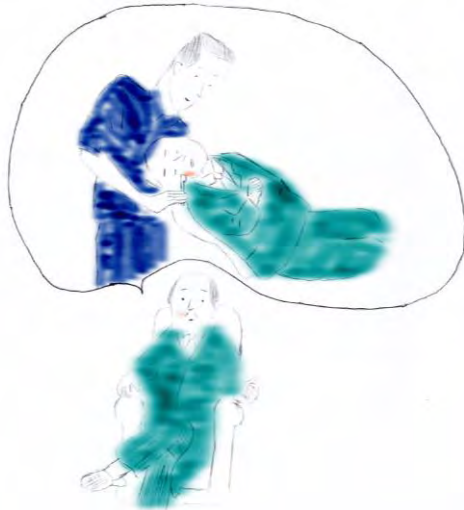
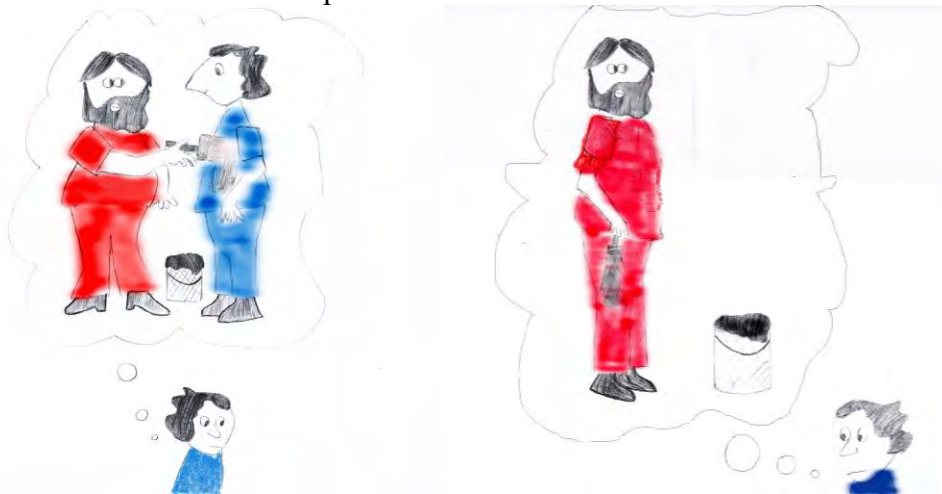2 Mr. Green said that Mr. Blue cut him. (Same pictures as above)
3. Mr. Green said angrily that Mr. Blue hit himself.

4. The man next to Mr. Green looked at himself in the mirror.

5. Mr. Blue said that Mr. Red painted himself.

B. Read each sentence, and then indicate who is performing the action by answering each question.

1. Lee remembered that Andrew introduced himself at the meeting.
Whom did Andrew introduce?
a. Lee b. Andrew c. somebody else d. either Lee or Andrew

2. Mary said that Jill understands herself.

Whom did Jack understand?
a. Mary b. Jill   c. either Mary or Jill d. somebody else
3. The hungry boy Jack was happy that John brought himself some cakes.
Whom did John bring cakes for?
a. Jack b. John c. either Jack or John d. somebody else
4. Mrs. Brown was angry that Mrs. Black blamed her for the wrong decision.
Whom did Mrs. Black blame?
a. Mrs. Brown b. Mrs. Black c. either Mrs. Brown or Mrs. Black d. somebody else
5. The lady beside June hit herself.
Whom did the lady hit?
a. the lady b. June c. Either the lady or June d. somebody else
6. Bill, who helped Tom, blamed himself.
Whom did Bill blame?
a. Bill b. Tom c. either Bill or Tom d. somebody else
7. The depressed woman said that the mother should leave her alone.
Who should be left alone?
a. the woman b. the mother c. either the woman or the mother d. somebody else.

C. Select a female classmate or relative or friend of yours (mother, sister, aunt, niece, etc.) and write her name below. Which of the following statements is likely true to her? Name_____
1. She says that her father will buy himself a new car. True____   Not True___
2. She hopes that her father would buy himself a house in Hawaii. True___   Not True___
3. She once said that her grandmother could take care of herself. True____   Not True___
4. Her grandmother once asked her to bring herself cookies to school. True____   Not True___
5. She says that if she can speak English well her mother will be proud of herself. True____ Not True___

Select a male classmate or relative or friend and do the same.
Name_____
1 He says that his father will buy himself a new car. True____   Not True___
2. He hopes that his mother would buy herself a house in Hawaii. True___   Not True___
3. He once said that his grandmother could take care of herself. True____   Not True___
4. His grandmother once asked him to bring himself cookies to school. True____   Not True___
5. He says that if he can speak English well his father will be proud of himself. True____   Not True___

D. Select the best way to complete each sentence.
1. The little girl was happy that her father bought _____a nice toy.
a. herself b. himself c. either a or b
2. The cashier was angry that the woman wrote _____ a fake check.
a. himself b. herself c. either a or b
3. The waitress behind the man calmed ____ with a glass of wine.
a. himself   b. herself c. either a or b
4. The little girl in front of the policeman protected ____ from the poisonous plant.
a. himself   b. herself c. either a or b
5. The famous actor heard that the popular hostess talked about _____ on TV last week.

a. himself   b. herself c. either a or b

E. Listen to each sentence, and then indicate who is performing the action by answering each question.

1. The hungry housewife was grateful that the husband cooked himself during the weekend.
Who is "himself" in the statement?
a. the housewife b. the husband c. either the housewife or the husband d. somebody else

2. Mr. Brown said that Mr. Green sprayed himself with insect repellent.
Who is "himself" in the statement?
   a. Mr. Brown b. Mr. Green c. either Mr. Brown or Mr. Green d. somebody else

3. Mr. Black was angry that Mr. Green did not trust him.
Who is "him"?
   a. Mr. Black b. Mr. Green c. either Mr. Black or Mr. Green d. somebody else

4. Mr. Black knew that Mr. Brown killed himself.
Who is "himself"?
   a. Mr. Black b. Mr. Green c. either Mr. Black or Mr. Green d. somebody else

5. Mrs. Black was outraged that Mrs. Green locked herself in the car.
Who is "herself"?
a. Mrs. Black b. Mrs. Green c. either Mrs. Black or Mrs. Green d. somebody else

6. Mr. Brown dreamed that Mr. Green shot himself.
Who is "himself"?
a. Mr. Brown b. Mr. Green c. either Mr. Brown or Mr. Green   d. somebody else

F. Read the story and then complete the sentence according to the content of the story.

1.   Bill is a student. There was a new teacher in his class today. During class, the teacher asked Bill's name and Bill's hometown. Bill told the teacher that he was born in New York.
The teacher said that Bill introduced _____.

2.   Susan wanted a job in the hospital. A nurse interviewed Susan for the job. The nurse asked Susan about her experience, her education and whether she got on well with people. The nurse told the director about the interview.
The nurse said that Susan talked about _____ in the interview.

3.   A young boy was looking at one of Mr. Robins's guns. The young boy accidentally pulled the trigger and the gun fired. Unfortunately, the bullet hit Mr. Robins in the arm.
Mr. Robins realized that the boy shot _____ accidentally.

4.   Annie wanted to get her friend into trouble. Annie went to the closet and suggested that her friend lock the door from the outside. When her friend followed her advice, Annie started to cry. When Annie got out finally, she told her mother about it.
Annie said that her friend locked _____ in the closet.

G . Read the passage. Then answer the questions that follow.

Parent: Can my son drive (1) *himself* to games, Director?

Director: The university provides transportation to all games and students are expected to travel together as a team. Under special circumstances, the team director may allow a student to drive (2) *himself* only after a parent has made those arrangements with the director in advance.

Student: The team director says that a student cannot drive *(3) himself* to games except under special circumstances.

1. Who does the underlined (1) *himself* refer to?
a. director b. son c. either a or b d. somebody else.

2. Who does the underlined (2) *himself* refer to?

a. the team director b. a student c. either a or b d. somebody else.
3. Who does the underlined (3) *himself* refer to?
a. the team director b. a student c. either a or b d. somebody else.
4. In the underlined (3) *himself,* can we replace himself with herself?
Yes __ No___

Appendix D: Results of Reaction Times

		Pretest			Posttest		
		t	*df*	Sig.	*t*	*df*	Sig.
Subject analysis							
Neutral sentences:							
Both match vs. no match:							
Position 1	BU1–NU1	-.44	27	.332	.16	27	.438
Position 2	BU2–NU2	-.75	27	.229	-.58	27	.284
Position 3	BU3–NU3	-2.0	27	.030*	-2.5	27	.011*
Position 4	BU4–NU4	-1.6	27	.062	-.39	27	.351
Only NP1 match vs. only NP2 match:							
Position 1	OU1– TU1	.05	27	.481	1.3	27	.104
Position 2	OU2– TU2	.31	27	.379	-.82	27	.212
Position 3	OU3–TU3	1.6	27	.057	2.4	27	.011*
Position 4	OU4–TU4	-.44	27	.332	1.6	27	.065
Biased sentences:							
Both match vs. no match:							
Position 1	BG1–NG1	1.7	27	.051*	.19	27	.425
Position 2	BG2–NG2	-.67	27	.254	-1.2	27	.124
Position 3	BG3–NG3	-2.2	27	.018*	-3.4	27	.001*
Position 4	BG4– NG4	.13	27	.451	-3.4	27	.001*
Only NP1 match vs. only NP2 match:							
Position 1	OG1– TG1	.79	27	.220	.11	27	.459
Position 2	OG2–TG2	-.51	27	.309	-.21	27	.417
Position 3	OG3– TG3	-.78	27	.222	.55	27	.293
Position 4	OG4–TG4	-.22	27	.414	1.9	27	.032*
Item analysis							
Neutral sentences:							
Both match vs. No match:							
Position 1	BU1–NU1	-.20	47	.421	.08	47	.468
Position 2	BU2–NU2	-.34	47	.368	.36	47	.360
Position 3	BU3–NU3	-2.0	47	.028 *	-1.9	47	.031*
Position 4	BU4–NU4	-1.1	47	.138	-.16	47	.438
Only NP1 match vs. only NP2 match:							
Position 1	OU1– TU1	.02	47	.491	.84	47	.203
Position 2	OU2– TU2	.83	47	.207	-.73	47	.235

Position 3	OU3–TU3	1.2	47	.127	2.0	47	.025*
Position 4	OU4–TU4	-.24	47	.406	1.3	47	.104

Biased sentences:

Both match vs. No match:

Position 1	BG1–NG1	1.4	47	.083	-.12	47	.451
Position 2	BG2–NG2	.06	47	.475	.14	47	.443
Position 3	BG3–NG3	-1.6	47	.054*	-3.0	47	.003*
Position 4	BG4– NG4	-.15	47	.441	-3.3	47	.001*

Only NP1 match vs. only NP2 match:

Position 1	OG1– TG1	.54	47	.298	.53	47	.301
Position 2	OG2–TG2	-.23	47	.409	-.54	47	.297
Position 3	OG3– TG3	-.65	47	.261	.32	47	.375
Position 4	OG4–TG4	-.14	47	.446	1.5	47	.067

Note. The first letter indicates type of conditions (B = both match; N = no match; O=only NP1 match; T=only NP2 match); the second letter indicates sentence categories (U = neutral; G = biased); the number indicates the test position. Thus, BU1−NU1 means the comparison of RT means on the "both match" and "no match" sentences involving the structural cue at the first position.