The Effect of Input-based Instruction Type on the Acquisition of Spanish Accusative Clitics

Justin White

Hispania 98.2 (2015): 264–84

Hispania Open Access files are licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
The Effect of Input-based Instruction Type on the Acquisition of Spanish Accusative Clitics

Justin White
Florida Atlantic University

Abstract: The purpose of this paper is to compare structured input (SI) with other input-based instructional treatments. The input-based instructional types include: input flood (IF), text enhancement (TE), SI activities, and focused input (FI; SI without implicit negative feedback). Participants included 145 adult learners enrolled in an intermediate Spanish course. Learners completed a pretest, treatment, immediate posttest, delayed post-test (three weeks after treatment) and an extended delayed post-test (six weeks after treatment) measuring both interpretation and production of the Spanish accusative clitics. The results revealed significant findings for all groups on the interpretation task, with SI being the most robust. The findings for the production measures are not so straightforward; however, all groups showed some improvement. The theoretical and methodological ramifications of these findings are discussed.

Keywords: input enhancement/realce de input, pronouns/pronombres, second language acquisition/adquisición de segundas lenguas, Spanish/español, structured input/input estructurado

Given that all major theoretical frameworks in second language acquisition (SLA) posit a fundamental role for input (e.g., Ellis 2007; Gass and Mackey 2007; VanPatten 2007; White 2007), instructed SLA research has begun investigating ways in which instruction can enhance L2 learners’ input processing through exposure to input via various pedagogical interventions. In fact, Doughty (2003) has argued, “the goal of L2 instruction should be to organize the processing space to enable adults to notice the cues located in the input, as they did when they were children. A guiding principle in this regard is to engage perceptual processes during implicit learning, rather than processes that promote metalinguistic awareness” (298). In this light, the issue is not whether input is necessary, but rather whether all input is equally effective or if some forms of input are more beneficial than others. Therefore, the present study investigates various input-based instructional interventions and measures their effects on acquisition of the Spanish accusative clitics.

Research on exposure to input suggests that learners’ default processing strategies may be non-optimal and thereby inhibit their correct processing of some language features and structures. Learners, at least in beginning stages of acquisition, tend to tag the first noun in a noun phrase-verb-noun phrase (NP-V-NP) sequence as the subject/agent in the sentence when it actually may not be the subject/agent. For example, La ve Juan (lit: Her ACC sees Juan NOM) may be misinterpreted as ‘She sees Juan.’ instead of ‘Juan sees her.’ This incorrect processing is especially problematic when the object is assigned the role of subject, when in fact it is the receiver of the action, and consequently causes erroneous interpretation of the input. The issue at hand goes beyond the learner just getting the word order wrong, rather, the consequences for this type of misinterpretation for language learners is that the acquisition of some structures, such
as passives, object-verb-subject OVS structures, case marking, and so forth, may be delayed.
Providing the most comprehensive account of input processing, VanPatten (2007) describes
this tendency as the first noun principle.

What is clear is that left to their own devices, learners may process the input incorrectly,
and therefore, need something else to aid them in correctly parsing sentences and in making
the appropriate form-meaning connections. What exactly learners need to facilitate their correct
processing of the input is a focus of the current study. Therefore, the study at hand investigates
the effects on acquisition of the Spanish third accusative clitics of exposure to various input-
based instructional treatments: structured input (SI), a variation of SI called focused input (FI)
in which the inherent variable of implicit negative feedback in SI is isolated, input flood (IF),
and input flood with text enhancement (IFTE).1

Background

Processing Instruction (PI) and Structured Input (SI)

Among the various options of input-based instructional interventions, processing instruc-
tion (PI) and one of its components, SI, have been the most widely researched. PI is an input-
based pedagogical intervention stemming from one model of input processing (VanPatten 1996,
2004, 2007) that targets the non-optimal default processing strategies learners employ while
processing input in order to make form-meaning connections. This pedagogical intervention
consists of three main components: explicit information about the target structure, its related
processing problem, and SI activities. SI activities are a specific type of input-based activity
designed to alter learners’ incorrect processing strategies and help them make the appropriate
form-meaning connections by attending to the target form for meaning.

What distinguishes PI/SI from other types of pedagogical interventions is that they are
directed at altering learners’ incorrect processing strategies as identified by the principles of
input processing (VanPatten 2007). Processing strategy training contrasts with other types
of input enhancement techniques such as IF or TE, which are predicated on learners’ target form
noticing based on the noticing hypothesis. Schmidt (1990) claims that, “noticing is the necessary
and sufficient condition for the conversion of input to intake for learning” (17). According to
Schmidt (1990), without noticing, acquisition will not happen.

For two decades, the effectiveness of PI has been investigated with L2 learners of Span-
ish (e.g., Fernández 2008; Morgan-Short and Bowden 2006; VanPatten and Cadierno 1993;
VanPatten and Oikkenon 1996); French (Benati and Lee 2008; VanPatten and Wong 2004),
English (Benati and Lee 2008); Italian (Benati 2001, 2004); and Japanese (Lee and Benati 2007).
Detailed discussions of these studies, and others, can be found in Benati and Lee (2008) and
Lee and Benati (2007). Consistently, PI has been shown to be more effective than traditional
production-based instruction and that the results are attributable to SI activities and not to the
Explicit Information component of PI (e.g., VanPatten and Oikkenon 1996).

One limitation of PI/SI research to date is that it has consistently contained an element of
implicit negative feedback. As mentioned by Sanz and Morgan-Short (2004), during a certain
type of referential SI activity, the picture selection task, learners receive implicit negative feedback
when questions are answered correctly. During these referential SI activities, learners are given
two pictures with one accompanying sentence and asked to choose the picture that matches
up with the sentence. If they choose the correct picture, they are given a computer delivered
written response of ‘OK’ (or if conducted in class, the feedback is given orally), signaling that
they chose the correct match and then allowing them to continue on to the next question. The
implicit negative feedback presents itself at the point in which learners are informed of which
answer is correct and (by extension) which one is not. Figure 1 provides an example of the SI
picture selection task with the implicit negative feedback highlighted.
Sanz and Morgan-Short (2004) point out (as a limitation in their study), that because all participants received implicit feedback, they could not “determine whether it was the practice alone, or the practice in combination with implicit feedback, that equalized the performance on all four groups” (72). In order to address this question, it is necessary to isolate the variable of implicit negative feedback in the SI activities to determine if it is a contributing factor in the effectiveness of SI activities or if it is the exposure to the positive evidence in the input alone. The current study addresses this by isolating the implicit negative feedback variable in an input-based treatment called focused input FI.

To date, SI has not been compared with other input-based instructional interventions, with the exception of Morgan-Short and Bowden (2006), which included an exposure only group that received an input flood (referred to as the control group), in addition to a meaningful output-based instruction (MOBI) group, all of which targeted the Spanish accusative clitics. Their findings indicate no group performed significantly better than the other on either mode of the delayed posttest (interpretation or production), thereby suggesting that any of these instructional treatments, including the input flood, may be enough for learners to acquire this target form. These findings imply that the success of PI might be attributed to exposing learners to input alone.

Therefore, it is fitting to question whether the effectiveness of SI activities is attributable to the following: 1) learners simply receiving comprehensible input; 2) input in SI is structured in such a way that pushes learners away from non-optimal processing strategies; or 3) input in SI is coupled with the implicit negative feedback. The present study seeks to shed light on these questions by isolating the feedback component of SI activities and by comparing SI with other input-based treatments. The following sections address two other input-based instructional interventions included in the present study: input flood and Text Enhancement, which are both input enhancement techniques predicated on learners’ noticing the target forms in the input (as opposed to being predicated on learners’ processing strategies as with SI and FI).
Input Flood

Research has investigated the effects of manipulating learners’ exposure to input through types of input enhancement, a term coined by Sharwood-Smith (1993) to refer to any attempt to draw learners’ attention to a grammatical form while at the same time directing them to process for meaning. One such technique is input flood, which consists of providing learners with lots of instances of a particular target structure in oral or written form, thereby, saturating the input with the form to be acquired. The basis for this flooding is that learners will be more likely to notice and process linguistic items that frequently occur in the input (Gass 1997). During a typical IF task, learners are held responsible for the meaning of the text by means of content questions related to the text, and not asked any questions dealing specifically with the target forms embedded (flooded) in the text.

The findings of studies conducted in the area of IF are split; some studies such as Trahey and White (1993) and Spada and Lightbown (1999) suggest that IF may be effective for learners to understand what is possible in a language (i.e., SVO word order in English or question formation in English, respectively); however, it may not be an effective method for learners to acquire what is not possible in a language. On the other hand, the results of Williams and Evans (1998) suggest that the positive evidence in IF alone may be enough to facilitate acquisition of some forms (i.e., English passive forms) but not others (i.e., English participial adjectives).

Text Enhancement

Another input enhancement technique also predicated on learners’ target form noticing is text enhancement, which refers to typographical modifications (i.e., changes in font size, type, color, boldness) made to target forms in the input. Wong (2005) states that “this is essentially the idea behind textual enhancement: to render more salient particular features of written input that learners normally may not notice and make form-meaning connections for” (49). As with IF, learners are held responsible for the content or meaning of the text itself but they are not immediately quizzed on the enhanced target form.

Studies investigating the effects of TE have also produced mixed results; Shook (1994), Alanen (1995), Jourdenais, Ota, Stauffer, Boyson, and Doughty (1995), and White (1998) all found positive effects for TE when comparing groups exposed to either textually enhanced or textually unenhanced input; however, the results of White (1998) suggest that learners still need some other type of instruction to increase accuracy. Overstreet (1998), in fact, not only found that the enhanced text drew learners’ attention to the target forms, but he also noticed that there was a negative effect revealing that text comprehension decreased with the presence of the enhanced text. In this case, the enhanced text actually drew learners’ attention away from the meaning of the text, drawing it instead to the forms themselves. In a meta-analysis of sixteen studies, Lee and Huang (2008) found that learners exposed to enhanced texts outperformed those exposed to unenhanced texts by a minimal amount ($d = 22$); however, they advise caution in the interpretation of this figure because the effect size values were calculated by contrasting experimental groups with comparison groups instead of true control groups.

On the other hand, Izumi (2002) found no difference between enhanced or unenhanced input (paired with output). Similarly, Leow Egi, Nuevo, and Tsai (2003) found that some forms may be noticed more than others (either enhanced or unenhanced) but that enhancing the target forms does not produce differences in noticing or differences in performance on comprehension. In another study, Leow (2001) also showed a lack of effect for target form noticing.

The findings for the input enhancement techniques based on noticing (IF and TE), are not as consistent as those based on learners’ processing strategies (PI and SI). Notwithstanding, it is important to include these instructional techniques together in the same research design in
order to address whether the effectiveness of PI/SI is mainly due to learners being exposed to comprehensible input or whether it is specifically the manipulated input in PI/SI that is effective. The current study addresses one particular target form, Spanish accusative clitics, which has not been investigated in IF or TE research to date. Additionally, PI/SI have not previously been compared to other input-based treatments, such as IF or IFTE, and the present study addresses this limitation. This study also proposes to tease apart the variables of SI activities to determine if it is the input itself, the variable of implicit negative feedback, or the way in which the input is structured to push learners away from non-optimal processing strategies that is contributing to SI’s effectiveness. By investigating the effects of these input-based treatments, this study will shed light on whether input alone is enough to facilitate the necessary form-meaning connections, which will contribute to both PI/SI and input enhancement research, and ultimately have implications for language teaching.

Research Questions

Based on the studies reviewed in the previous sections, the following research questions guided the present study:

1) Do input-based instructional treatments lead to improved performance on L2 Spanish learners’ interpretation and production of third person accusative clitics?

2) If the input-based instructional treatments lead to improved performance, are the effects of the treatments comparable?

To this end, the present study includes the following four treatment groups: structured input (SI), focused input (FI), input flood (IF), and input flood with text enhancement (IFTE) and a control group receiving no instruction.

Method and Procedures

Participants

The participants in this study were drawn from nineteen sections of third semester language courses at Florida State University and each intact class was randomly assigned to one of the four treatment groups or the control group (who received no target form treatment and only received the same assessment measures). The initial participant pool consisted of 290 participants (N = 290) and a final N size after attrition of those who participated in all four test times of 145 ranging in age from 18 to 22 years old. The treatment group breakdown is the following: IF, (N = 24); IFTE, (N = 34); FI, (N = 32); SI, (N = 28); control group: (N = 27). Participants were deemed suitable for this study if (a) their native language was English, (b) they had no uncorrected visual impairment, (c) they were present for all testing and treatment sessions, and (d) scored lower than 71% (five out of seven questions correct) on the interpretation pretest.

Materials

Assessment Tasks

A split-block design was used that included four versions of the assessment measures described below: A, B, C, and D. The interpretation measure in the current study consisted of fifteen items, seven critical (OVS) and eight SVO sentence type distractors. Each question displayed two pictures and one sentence in written form. The written sentence matched up with one of the pictures, and the other picture represented the incorrect processing of the sentence. The pictures were labeled 'A' and 'B.' Students chose between pictures A, B, or an option labeled
‘Not sure.’ The instructions for participants indicated that they should select the picture that matched with the sentence they read, and if they were not sure, then they should choose the option ‘Not sure.’ Samples of the assessment measures are found in the appendix.

The sentence-level production task consisted of thirteen items: eight critical and five distracters. The questions consisted of a sequence of two pictures along with one incomplete sentence. The first part of the sentence described the first of the two pictures and the last part of the sentence was left uncompleted with a verb in parentheses (describing the second of the two pictures). The task required participants to finish the sentence according to what was depicted in picture number two. The goal of the task was to incorporate the accusative clitics to complete the sentences. Although it was possible to use nouns instead of accusative clitics to complete the sentences, the sentences more naturally lent themselves to the use of the clitics.

**Instructional Materials**

The target form investigated in this study was the Spanish accusative clitics in the preverbal form (OVS type sentences). The target forms for all treatment groups included six exemplars of *la* (‘her’), *lo* (‘him’), *las* (‘them,’ feminine), and *los* (‘them,’ masculine or masculine/feminine) for a total of 24 exemplars for each treatment. Sanz and Morgan-Short (2004) used 56 exemplars, more than double used in this study; however, they used the entire paradigm (first, second, and third person forms). Fernández (2008) used 30 SI items total in the third person singular and plural forms: 20 OVS and 10 SOV and SVO type sentences. At no point during treatment in any of the groups were learners required to produce the target form. In all cases, learners were only exposed to written input of the target forms. Samples of the instructional materials are also found in the appendix.

**Structured Input (SI)**

The materials consisted of twenty-four instances of Spanish third person accusative clitics located in twenty-four sentences in addition to six sentences (SVO) that established agential subject/object relationships similar to those with the target form but with nouns in place of accusative clitics. The SI treatment materials were based on those in Sanz and Morgan-Short (2004) and Fernández (2008), which were computer delivered adaptations of activities used in VanPatten and Cadierno (1993). The pictures in these materials were the same as those used in the aforementioned studies taken directly from Fernández’s treatment materials, making occasional modifications for picture clarity. Following Fernández (2008), only referential activities were used.

SI items consisted of two components: a short sentence and two pictures. Each sentence with a target form token contained the target item in either sentence initial position or sentence medial position. The picture pairs contained one image illustrating the event matching the sentence and one image illustrating the commonly interpreted incorrect meaning of the sentence. In other words, one of the pictures depicted the correct meaning of the sentence and the other picture illustrated the first noun being incorrectly processed as the subject, instead of as the object. In order to complete the activity, learners had to choose which picture they believed depicted the events in the sentence.

All sentences were similar in length, vocabulary level, and were controlled for probability of events. After the participants chose a picture that they thought illustrated the meaning of the sentence, they received feedback. Feedback in this study was operationalized by informing participants whether their answer was incorrect by the word ‘incorrect’ appearing written on the screen, or correct, by the word ‘correct’ appearing on the screen. In either event, participants moved directly on to the next item. Fernández (2008) also operationalized feedback in a similar way.
Focused Input (FI)

FI materials addressed the variable of implicit negative feedback inherent in SI activities while providing learners with positive evidence of uses of the accusative clitics. The FI activity contained one picture and one sentence, compared with SI activities that included two pictures and one sentence. In FI, participants received one picture with its corresponding sentence, with the target form (24 sentences, OVS and SOV types) or without the target form (six sentences, SVO type), and read the sentence that matched up with the picture. They were the same sentences used in the SI materials, the difference being that participants saw only one picture along with each sentence. Participants progressed (by clicking ‘next’ on their computer) from one example to another without any intervening information such as feedback. Learners received instructions that stated, “You are about to see a series of slides that contain a picture along with a corresponding sentence in Spanish. All the Spanish sentences correctly correspond with the picture. When you think you understand how the sentence explains the picture, press any button to move to the next slide.” Figure 2 contains an example.

Figure 2. An example of a focused input (FI) item (Translation: ThemACC follows PedroNOM ‘Pedro follows them’)

Input Flood (IF) and Input Flood with Text Enhancement (IFTE)

The IF and IFTE treatment materials consisted of four reading passages, each focusing on one of the third person accusative clitics. The IFTE materials were the same as those of the IF with one exception: the target forms were bolded and the font size was made larger (14 point font) than the surrounding text (12 point font). All of the passages contained between 125 and 133 words, 13 or 14 sentences, and the target forms were always in sentence medial location.

Participants in both the IF and IFTE groups read the passages at their own pace. Participants clicked ‘next’ when they finished reading the passage. After reading each story, participants answered three multiple-choice questions in English about the content of the passage. Each question appeared individually along with the passage on a separate page. There were four passages, three questions for each. One content question for two of the passages targeted the relationship between the protagonists of the passages, although it was not related to the subject/object relation established by the target form. During treatment, the passages were included
along with each question in order to account for exposure time to the target forms across treatment conditions.

Procedure

The pretreatment materials used in this study included a paper-and-pencil pretreatment packet including a consent form, a language history questionnaire, and an answer sheet for the pretest. The treatment materials for this study were delivered using the computer program SuperLab and accessed on computers in the computer laboratory on treatment day.

A week prior to treatment, participants who were recruited during class completed the language history questionnaire and a pretest (Time 1) measuring participants’ accuracy on interpretation and production of the Spanish third person accusative clitics. During recruitment, the instructor read a description of the investigation as dealing with Spanish grammar, informed students of the procedures of the study, and indicated that their participation was voluntary. They were also informed that by participating in the study they would receive a 100% on their lowest completed homework set (roughly 1% of their course grade). No one chose not to participate.

The pretest materials were projected onto the projector screen in the front of the classroom, and participants recorded their answers on a paper/pencil answer sheet. To be sure not to provide extra input, the researcher conducting the experiment had participants complete the production measure before completing the interpretation measure. The purpose of the pretest was to measure participants’ previous knowledge of the target form prior to treatment. To avoid a ceiling effect and for learners’ data to be included in the analyses, a criterion of a score below 71% was set on learners’ interpretation pretest scores. After completing the pretest, participants completed the background questionnaire.

The treatment lasted 50 minutes, and on treatment day students met in the computer laboratory in lieu of meeting in their classroom. The researcher (or assistant) provided all participants with an activity packet including procedures for how to access the materials on the computer and a paper/pencil answer sheet for posttest 1 (Time 2). Posttest 2 (Time 3) was conducted three weeks from the day of the treatment, and a posttest 3 (Time 4) was conducted six weeks after the treatment, both in the participants’ regular classroom and following a similar procedure to the pretest.

Scoring

For the interpretation task, participants received scores on a one-point scale (seven points possible). They received one point for answering the question correctly, zero points for answering the question incorrectly, and zero points for answering, ‘not sure.’ The production task was scored similarly in that a correct answer utilizing the correct accusative clitic merited one point and an incorrect or no answer at all was scored as a zero.

Results

Interpretation Task

Table 1 and Figure 3 present the descriptive statistics for the interpretation task. To ensure that there were no pre-existing differences among the treatment groups, the pretest scores for all groups were submitted to an ANOVA. The ANOVA revealed no main effect for group $F(4, 144) = 1.68, p = .166$. For this reason, it can be assumed that any gains in performance were due to instructional treatment type.

In order to determine whether the treatment type lead to increased performance on the interpretation task, separate ANOVAs with repeated measures were conducted for each of
Table 1. Descriptive statistics (in percentages) for treatment groups on the interpretation task

<table>
<thead>
<tr>
<th>Group</th>
<th>Time 1</th>
<th>Time 2</th>
<th>Time 3</th>
<th>Time 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input Flood (IF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>40.00</td>
<td>43.60</td>
<td>52.40</td>
<td>54.00</td>
</tr>
<tr>
<td>SD</td>
<td>9.13</td>
<td>22.15</td>
<td>21.66</td>
<td>22.9</td>
</tr>
<tr>
<td>IF + Text Enhancement (IFTE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>40.59</td>
<td>45.59</td>
<td>44.16</td>
<td>51.47</td>
</tr>
<tr>
<td>SD</td>
<td>13.01</td>
<td>20.33</td>
<td>20.76</td>
<td>17.78</td>
</tr>
<tr>
<td>Focused Input (FI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>36.88</td>
<td>61.25</td>
<td>45.94</td>
<td>58.75</td>
</tr>
<tr>
<td>SD</td>
<td>9.98</td>
<td>26.49</td>
<td>17.20</td>
<td>17.92</td>
</tr>
<tr>
<td>Structured Input (SI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>39.66</td>
<td>72.41</td>
<td>59.31</td>
<td>58.28</td>
</tr>
<tr>
<td>SD</td>
<td>10.52</td>
<td>20.99</td>
<td>26.04</td>
<td>19.83</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>44.14</td>
<td>36.55</td>
<td>42.07</td>
<td>42.76</td>
</tr>
<tr>
<td>SD</td>
<td>12.11</td>
<td>15.42</td>
<td>22.74</td>
<td>16.01</td>
</tr>
</tbody>
</table>

Figure 3. Group × test interaction on the interpretation test.

among each time. The significant findings are displayed in Table 2. The pairwise comparisons revealed the following findings. First, all input treatment types lead to increased performance between the Time 1 and Time 4, which was conducted six weeks after the treatment. Second, only two input treatment types, SI and FI, showed an improved performance between Time 1
and Time 2. FI also increased performance on Time 4 from Time 3. Finally, these same two groups showed a decrease in performance from Time 2 to Time 3. That said, both of these groups’ scores on Time 4 were still greater than Time 1.

The data for all interpretation tests were submitted to a $5 \times 4$ ANOVA with repeated measures. The between-groups variable was Group and the within-groups variable was Time. The repeated-measures ANOVA revealed a main effect for the within-subjects variable Time, $F(3, 432) = 17.08, p < .001, \eta_{\text{partial}}^2 = .106$; a main effect for the between subjects variable Group, $F(4, 144) = 7.81, p < .001, \eta_{\text{partial}}^2 = .178$; and a significant Time $\times$ Group interaction, $F(12, 432) = 5.63, p < .001, \eta_{\text{partial}}^2 = .135$.

To examine the main effect for Group, pairwise comparisons were conducted with a Bonferroni adjustment for multiple comparisons. These contrasts revealed the following: SI outperformed all groups except for FI ($p < .05$), and the FI group was the only group besides SI to outperform control ($p = .023$). No other significant contrasts were found. For the main effect for Time, pairwise comparisons revealed that all three posttests (immediate, delayed, second delayed) were greater than the pretest, $ps < .001$.

In order to explore the Group $\times$ Time interaction, the Bonferroni tests revealed the following differences among groups at each test: no difference among groups at the Time 1 ($p = .165$); SI and FI outperformed all other groups and the control at Time 2 ($ps < .05$), but there was no difference between FI and SI; SI outperformed control and IFTE at Time 3 ($p = .03$); and both SI and FI outperformed all treatment groups and the control at Time 4 ($ps < .05$), but there were no differences between SI and FI. IF and IFTE did not significantly outperform the control group at any time.

In general, the findings for the interpretation test revealed that all groups, except the control group, made significant gains over time in the correct interpretation of Spanish third person accusative clitics. In terms of differences among the treatment groups, the findings for SI were the most consistent. SI was the only treatment group that outperformed the control on all of the

| Table 2. Pairwise comparisons between interpretation tests by treatment group |
|------------------|------------------|------------------|
| Group            | Mean Difference  | $p$              |
| Input Flood (IF) |                  |                  |
| Time 4 > Time 1  | 14.00            | .031             |
| Time 4 > Time 2  | 10.40            | .006             |
| IF + Text Enhancement (IFTE) |          |                  |
| Time 4 > Time 1  | 10.88            | .012             |
| Focused Input (FI) |                 |                  |
| Time 2 > Time 1  | 24.38            | <.001            |
| Time 2 > Time 3  | 15.31            | .019             |
| Time 4 > Time 1  | 21.88            | <0.001           |
| Time 4 > Time 3  | 12.81            | .010             |
| Structured Input (SI) |              |                  |
| Time 2 > Time 1  | 32.76            | <.001            |
| Time 2 > Time 3  | 13.13            | .019             |
| Time 4 > Time 1  | 18.62            | .001             |
In addition, SI also outperformed both input flood groups (IF and IFTE) on at least two of the posttests. It is also worthwhile noting that FI was the only other treatment group to outperform the control on at least one posttest. Furthermore, there were no significant differences between FI and SI on any of the posttests.

Sentence-level Production Task

Table 3 presents the descriptive statistics for the treatment groups on the sentence-level production tests. As the data in the table indicate, the scores are quite low. For this reason, to determine whether the various treatments led to increased performance on production over time, five separate Friedman tests were conducted for each group. These nonparametric tests were conducted because of the non-normal distribution of the data, and the basic assumptions of the ANOVA could not be met. The alpha used throughout was .05. The results of the Friedman tests revealed that the following groups demonstrated a significant increase in performance: IF, $\chi^2 (3, N = 24) = 9.37, p = .025$; IFTE, $\chi^2 (3, N = 34) = 7.01, p = .017$; FI, $\chi^2 (3, N = 32) = 16.11, p = .001$; and SI, $\chi^2 (3, N = 28) = 17.65, p = .001$. There was no significant finding for the control group: $\chi^2 (3, N = 27) = 3.63, p = .305$.

<table>
<thead>
<tr>
<th>Group</th>
<th>Time 1</th>
<th>Time 2</th>
<th>Time 3</th>
<th>Time 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input Flood (IF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>0.33</td>
<td>1.42</td>
<td>1.83</td>
<td>2.25</td>
</tr>
<tr>
<td>$SD$</td>
<td>1.63</td>
<td>3.45</td>
<td>4.50</td>
<td>4.54</td>
</tr>
<tr>
<td>IF + Text Enhancement (IFTE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>0.50</td>
<td>2.74</td>
<td>2.74</td>
<td>1.82</td>
</tr>
<tr>
<td>$SD$</td>
<td>1.80</td>
<td>5.65</td>
<td>5.65</td>
<td>3.05</td>
</tr>
<tr>
<td>Focused Input (FI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>0.00</td>
<td>2.90</td>
<td>1.44</td>
<td>1.69</td>
</tr>
<tr>
<td>$SD$</td>
<td>0.00</td>
<td>6.79</td>
<td>2.87</td>
<td>4.40</td>
</tr>
<tr>
<td>Structured Input (SI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>0.18</td>
<td>1.89</td>
<td>3.21</td>
<td>2.57</td>
</tr>
<tr>
<td>$SD$</td>
<td>0.94</td>
<td>4.92</td>
<td>4.95</td>
<td>3.41</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>0.00</td>
<td>0.11</td>
<td>0.52</td>
<td>0.26</td>
</tr>
<tr>
<td>$SD$</td>
<td>0.00</td>
<td>0.58</td>
<td>1.95</td>
<td>1.02</td>
</tr>
</tbody>
</table>

In order to determine which groups improved on which tests, pairwise comparisons with a Bonferroni adjustment for multiple comparisons were conducted. The significant findings are revealed in Table 4. The pairwise comparisons revealed the following major findings. IF lead to increased performance from Time 1 to Time 4, IFTE from Time 1 and Time 3, FI from Time 1 to Times 2 and 4, and SI from Time 1 to Times 3 and 4.

In order to test whether any differences existed among the treatment groups, the production scores for the three posttests were determined for each input group (IF, IFTE, FI, and SI), and the data were submitted to a Kruskal-Wallis test with Group as the between-subjects variable. The results revealed no significant differences among the four input groups, $\chi^2 (3) < 6, ps > .10$. 
This study set out to examine the effects of the input-based instruction types (input flood [IF], input flood with text enhancement [IFTE], structured input [SI], and focused input [FI]) on interpretation and production of the Spanish third person accusative clitics. Instead of comparing the effectiveness of SI with other production-based activities, as in previous research, this study focused exclusively on comparing SI with other input-based instructional treatments.

In order to explain the findings of this study, I will return to the original research questions.

Research question 1) Do input-based instructional treatments lead to improved performance of the target form? The answer to the first research question is ‘yes.’ All groups (except control) showed significant improvement on the interpretation and production measures. These findings suggest that, for this form, input is enough to improve learners’ interpretation and production of accusative clitics.

These findings, however, differ slightly from those of Morgan-Short and Bowden (2006) with respect to the control group’s performance. In their study the control group showed improved
Hispania 98 June 2015

performance over time; however, the control group in their study was not a typical control group. That is, their control group actually received a sort of input flood of the target form as opposed to a typical control group receiving no input of the target form at all. The findings suggest that their exposure to the target form may have led to their improved performance, thereby explaining why they showed improved performance but the control group in this study did not. Therefore, a general implication from their study is that input may be enough to facilitate acquisition. In the case of the present study, even though all groups showed improvement, were there differences in performance among the groups?

This leads to the second research question in this study: are the effects of the treatments comparable? The answer to this question is twofold: the results of the treatments are not comparable for interpretation but in some ways are for production. For interpretation, even though all groups showed improvement, there were significant differences among some of the groups. Both the IF and IFTE groups showed improvement over time on interpretation measures; however, neither group performed significantly better than the control group on any posttest. Although it is likely that learners may have noticed the target forms while reading for meaning, as Izumi (2002) points out, noticing the target forms does not necessarily encourage the cognitive processes necessary for acquisition to take place. Therefore, the among group findings of this study suggest that IF (and IFTE) are no more effective than no instruction at all considering their comparable performance with that of the control group.

Regarding the FI and SI groups, no significant differences emerged between these two groups on any of the interpretation posttests. That said, only SI outperformed the IF and IFTE groups on two posttest times. The first question asked here is, what explains the similarities in the findings for SI and FI? One of the possible explanations for SI not outperforming FI is that the type of comprehensible input provided by these two input-based treatment types is similar enough to push learners to make form-meaning connections. The input in both SI and FI is manipulated so that learners are pushed to alter their incorrect processing strategy. The implicit negative feedback variable inherent in SI activities does not seem to be the driving force behind SI’s effectiveness but rather the type of input directed at altering learners’ incorrect processing strategies. In other words, both treatment groups’ materials contained input in which the first noun was not always the subject (e.g., *Lo ve María.* ‘Maria sees him’) and asked learners to select among two pictures to indicate the correct interpretation (SI) or to understand how one picture illustrates the meaning of the sentence (FI). Although all of the input-based treatments (IF, IFTE, FI, and SI) were meaning based treatments, only the SI and FI treatments pushed learners to pay attention to formal properties of Spanish in order to understand meaning (i.e., agent-patient relationships). Furthermore, only SI and FI are predicated on non-optimal processing strategies and attempt to push learners away from them. Perhaps, for this reason, these two treatment groups performed similarly.

Although no significant differences emerged between SI and FI, the findings for SI seem to be more consistent given that only SI significantly outperformed at least one of the input flood groups at each posttest time. Why might this be the case? Even though FI is a variation of an SI activity (SI without implicit negative feedback), one possible explanation is that learners are not required to react to the input in the same way as they do with the SI activities. In SI activities, there is an element of task essentialness; that is, learners must choose between two pictures to indicate how they interpret the sentence, whereas with FI, learners were provided one sentence that matched up with one picture as they advanced through the examples by clicking on the button box. Because the written input itself was the same for both groups, it may be that in SI activities learners are held accountable for their responses due to the nature of the difference between treatment types. That is, learners exposed to SI are forced to make a choice and are then informed of the correctness of that choice. This line of reasoning is consistent with recent proposals that acquisition is indeed an “error” or “failure-driven process” (e.g., Carroll 2007). That is, a
change in parsing procedures can only occur when the procedures or strategies in place to parse the input stream fail (Carroll 2001, 2007; VanPatten 2007). Perhaps then, learners completing the SI activities had opportunities to ‘fail’ in their comprehension during the activities. It could be precisely this failure and resulting implicit negative evidence that is aiding learners in readjusting their processing strategy to a more optimal one and pushing them to interpret the input correctly. Therefore, due to the nature of SI activities, the inherent element of implicit negative feedback along with the input is facilitating acquisition. Clearly more research in this area needs to be performed; however, these are possible explanations for the differences in performance among groups in the interpretation task.

On the production measure, all treatment groups (except control) showed improvement from the time of the pretest to posttests; however, no significant differences were found among any of the groups. Also, the findings for production are not as clear as those of the interpretation task due to the low scores. In general, we expect the interpretation scores to be more robust than the production scores given the nature of the two processes and the training involved. Learners, in all of the treatments in this study, were exposed to input in some form or another and did not produce the target form at any point during the treatment. Secondly, production requires a different set of processes than interpretation, and given that learners were trained to interpret the target form, effects on production will be lower than interpretation, as they have in previous studies. Additionally, on the interpretation task, learners have a choice of answers, whereas in the production task, learners cannot rely on ‘chance’ for a correct answer; thereby, making these two tasks intrinsically more difficult. However, in this case, the results on the production measure in the present study differ from those of previous studies. In VanPatten and Cadierno (1993), the range of gains from pretest to posttest was from 4.3 to 6.7. Sanz and Morgan-Short (2004) report mean gains ranging from 31% to 70%, and VanPatten and Oikkenon (1996) report means of .55 to 3.00 (out of 10), whereas in this study mean score increases ranged from 4% to 10%. In comparison to other studies, learners’ scores on the production measures were low. Why might this be the case?

One possible explanation for the low scores on both production measures in comparison with previous research is the order in which these tasks were completed. In other studies, learners completed the interpretation measure first, and then they completed the production measures. In such an event, the interpretation test provided input to the learner in which the target form was required to interpret the meaning of the sentence. In this study, however, learners first completed the production measure followed by the interpretation measure.

The ordering of the tasks was essential in order to control for the amount and type of input to which the treatment groups were exposed. The production measure was completed first so that learners were not exposed to more input of the target form through the interpretation measure. In this event, after learners completed the production task, they received the input from the interpretation task, but not until the last measure of the test. In this case, learners were not able to use the input from the interpretation measure (as they could in other studies) in order to complete the production task. Therefore, in this study, learners may not have even been aware of what the target structure was when completing the tasks. They relied purely on their exposure to the instructional treatment. Future research will have to investigate the effects of ordering of the assessment measures.

Another possible explanation is the amount of time that learners had to complete the task. In this study, the tests were displayed on the projector screen (with the exception of Time 2, which was on each learner’s individual computer monitor) using the Microsoft PowerPoint program, and each test question was automatically timed to move on to the next question. In this study, each question in the production measure was displayed for ten seconds; however, in Sanz and Morgan-Short (2004) learners were given twenty seconds to complete the same task. Ten seconds was chosen as an appropriate amount of time based on a pilot study demonstrating with learners
of a comparable proficiency level that ten seconds was enough time, yet not too long in order to minimize learners engaging in monitoring, thereby drawing upon metalinguistic declarative knowledge (Doughty 2003). However, during treatment, the researcher and assistants, both noticed that learners seemed “rushed.” Further testing will have to investigate how much time is needed for learners to complete this measure.

Some investigations have used instructor-led in-class treatments via instructor delivery (i.e., VanPatten and Cadierno 1993; VanPatten and Oikennon 1996) whereas others have used computer assisted delivery methods (i.e., Fernández 2008; Morgan-Short and Bowden 2006; Sanz and Morgan-Short 2004). Future research will have to investigate the effects of these treatment types delivered in both lab based and classroom based settings in order to determine if the delivery method plays a role in the findings.

A final possibility is related to the number of tokens in the study. The current study isolated the third person accusative clitics and included twenty-four target form examples (six examples of masculine singular, masculine plural, feminine singular, and feminine plural). Twenty-four tokens were used throughout this study considering that using that more tokens would have resulted in spreading the treatment over two class sessions, which, due to logistical planning issues, was not possible. This is contrasted with previous research in which treatment exposed learners to roughly five times the input in this study. For example, VanPatten and Cadierno (1993) included over a hundred tokens from the entire paradigm during instruction alone (not including the added input in test order). It is suspected, then, that the lower number of tokens in the current study may be contributing to the differences in results. Future research will have to investigate the issue of token frequency across treatment types.

Clearly, acquisition is a long process, and after a short training phase, any improvement can be seen as making steps in the right direction toward acquiring the target form. Even though mean scores for all groups did not exceed 10% on any of the production test times, two important points are to be made: 1) all treatment groups made significant improvement on production of the target form; and 2) participants’ exposure time to the target form was only limited to twenty-four instances of the target items during one treatment session. In this light, it is surprising that learners were able to produce anything at all after this brief of a treatment. Seeing any increase in performance based on such limited exposure, especially without explicit information about the form, is evidence that the treatment is doing something. Lee (2002) states, “acquisition starts with processing and occurs incrementally” (74). Language acquisition is a long, gradual process and does not happen overnight. Some grammatical forms take years to acquire, and in some cases learners do not reach native-like competency with them. In the case of the current study, considering that learners’ exposure to the target items was limited to one treatment session, and some of their starting points were zero, they are off to a respectable start. An anonymous reviewer also suggests that the low robustness of the findings on the production measure could be due to the very short duration of the treatment (50 minutes). Future research will have to investigate the effects of treatments spread over various days and compare the findings.

Conclusion

Some pedagogical implications can be taken from the findings of the current study. First, SI is an effective and plausible input type at pushing learners to make connections between form and meaning with this target form. One of the advantages to SI, (albeit true of all of the input-based instructional treatments in this study) is that it can be used via various delivery methods; classroom or via computer-assisted language learning tools delivered online or offline. In effect, there are many options for the delivery of SI and platforms that can support this type of activity with both aural and written input. One of the reasons why SI is well suited for in-class or online delivery, which is unique among the other input-based treatments, is due to the element of
‘forced choice’ inherent in this type of activity. Learners are forced to make a choice, in this case between two pictures of which one corresponds with a sentence, and they then receive feedback based on their response. There are also many options for feedback such as, but not limited to, written/aural, correct/incorrect, and explicit grammar oriented feedback. Due to these reasons, the incorporation of SI activities into textbooks and online activity components on publishing company platforms is both plausible and beneficial to students. For example, SI activities can be automatically graded, and they can include immediate and individual feedback.

Second, although effective, FI may not be the best pedagogical tool. First off, due to the nature of the activity learners are able to passively complete the task. Because of this, learners may also not see much of a point to the activity. It may be possible to use some examples of FI as a quiz or sort of game in class; however, solely as a pedagogical learning tool, it may not be the preferred use of time. This is mainly because of the affective effect on learners; they may just not get the point of the activity. After all, learners only receive positive evidence of the correct answer, and they do not receive feedback of any sort. However, as an anonymous reviewer points out, post-treatment interviews or focus groups may be necessary to gauge learners’ perceptions of this activity type.

Lastly, both of the input flood types (IF and IFTE), based on the findings of this study, may be overloading the cognitive processing constraints of the learner to the extent that learners are not able to attend to the target form. Because of the flood of input, learners may only focus on the meaning of the input alone and not the target forms. Pedagogically, learners may benefit from other aspects of reading passages such as vocabulary exposure, reading comprehension practice, pronunciation; however, for acquiring this target form, there may be better options, such as SI, at least for this level of learner. IF and IFTE types can also be included as part of an online component in a course and the content questions could be automatically graded, however, as the results of this study suggest, there are better options for this particular target structure.

Keeping in mind the limitations, the findings of this study suggest that structured input, a component of processing instruction, is not only effective when compared with production-oriented instructional techniques, but it is also effective when compared with other input-based treatments, thus shedding some light on why this kind of instruction is effective. It is hoped that this study will serve as a springboard to further research on a variety of processing problems as well as to provide a clearer picture of the role of these instruction types in promoting learners’ accurate production.

NOTE

1 The findings reported here are based on my doctoral dissertation submitted to Florida State University in 2008. I am deeply grateful to Michael Leeser for his guidance in my dissertation as director and chairperson of my committee, and to my other committee members. Many thanks are also due to Bill VanPatten for his guidance in the design of the study as well as his continued support.

WORKS CITED


**APPENDIX**

**Sample Materials**

I. Structured Input

1. [Participant reads: “Lo llaman sus padres por teléfono”] (“His parents call him”)

2. [Participant reads: “Las invita al cine Manuel”] (“Manuel invites them to the movies”)

---

281
II. Input Flood

El Restaurante
Dos mujeres con mucho dinero caminan por la calle. Un ladrón las ve y piensa, “ajá, voy a robar todo su dinero.” El ladrón las sigue a un restaurante. Las mujeres entran al restaurante. El mesero las saluda y van a una mesa frente a la ventana. El ladrón las escucha cuando están en la mesa. Las mujeres hacen su plan de ataque y deciden que van a robar el restaurante. El ladrón está sorprendido. De repente, las mujeres empiezan a gritar, “¡Vamos a robar todo su dinero!” El gerente abre la caja fuerte y ellas toman todo el dinero. El ladrón las mira. La gente grita y una persona llama a la policía. Llega la policía muy rápidamente y las llevan a la cárcel.

The Restaurant
Two women with a lot of money are walking down the street. A thief sees them and thinks, “Aha, I am going to steal all of their money.” The thief follows them to a restaurant. The women enter the restaurant. The waiter greets them and they go to a table in front of the window. The thief listens to them when they are at their table. The women make their plan of attack and decide they are going to rob the restaurant. The thief is surprised. All of a sudden, the women begin to scream, “We are going to steal all of your money!” The manager opens up the lock box and they take all the money. The thief watches them. The people scream and one person calls the police. The police arrive very quickly and they take them to jail.

1. Where do the women go?
   A. To their house
   B. To eat dinner
   C. To the bathroom

2. Who does the money come from?
   A. The women
   B. The waiter
   C. The manager

3. Who robs the restaurant?
   A. The bus boy
   B. The original thief
   C. The women

III. FI Materials

1. [Participant reads: “Las ve el niño”] (“The boy sees them”)
2. [Participant reads: “Lo escucha la abuela”] (“The grandmother listens to the boy”)

IV. Interpretation Task
Select the correct picture for each description and mark either a or b. If you are not sure, choose c.

1. [Participant reads: “Los saluda la mujer”] (“The woman greets them”)

   A
   B
   Not sure

2. [Participant reads: “La busca el niño”] (“The boy looks for her”)

   A
   B
   Not sure
V. Production Task

1. La policía reconoce a los ladrones y después ____________ (seguir).
   (The police recognize the thieves and then ____________ (follow)).

2. María encuentra a un amigo y luego ____________ (saludar).
   (María runs into a friend and then ____________ (greets)).