ATTENTION WHEN?

An Investigation of the Ordering Effect of Input and Interaction

Susan M. Gass and María José Alvarez Torres

Michigan State University

This paper investigates the effects of input and interaction as separate entities and in combination. We further investigate these effects as a function of different language areas. One hundred two learners of L2 Spanish were provided with input on (a) Spanish gender agreement (noun + adjective), (b) estar + location, and (c) seven vocabulary items. There were four conditions: (a) material focused solely on input, (b) material focused solely on interaction, (c) input-focused material followed by interaction, and (d) interaction-focused material followed by input. A control group completed a pretest and posttest. In general, greatest improvement from pretest to posttest for all conditions was noted for vocabulary. Learners exposed to input and interaction in combination showed greater improvement than those in conditions with only input or only interaction. In the two grammatical areas (gender agreement and estar + location), learners who received interaction followed by input showed greatest improvement. We consider issues such as complexity and abstractness to account for the findings of differential effects on language areas.

INPUT AND INTERACTION

It is widely recognized that input is essential for language acquisition. In addition to input, it is also accepted that interaction plays a crucial role in the process of learning a second language (L2). What has not been explored to

Funding for this project was provided by a federal grant from the U.S. Department of Education to the Center for Language Education and Research at Michigan State University (P229A990012 and P229A020001). We would like to thank George Sirbu and Pingping Ni for help with the statistics of this study. We are also grateful to the anonymous SSLA reviewers for helpful comments on an earlier version of this paper. All errors remain our own.

Address for correspondence: Susan Gass, English Language Center, A-714 Wells Hall, Michigan State University, E. Lansing, MI 48824; e-mail: gass@msu.edu.

© 2005 Cambridge University Press 0272-2631/05 $12.00
any great extent is the relationship between these two contextual factors of acquisition. This paper attempts to address this issue by exploring the order and consequent effect of the input-interaction relationship.

Since the early 1980s, the roles of input and interaction have been recognized as important in our understanding of how L2s are learned. In its simplest form, input is the sine qua non of acquisition. There is no theory or approach to SLA that does not recognize the importance of input, although theories differ as to its significance. For example, Ellis (2002) argued for the importance of frequency in SLA, a position that clearly relies heavily on the input to learners who, in his model, extract frequency patterns from that input in their attempt to learn an L2. Similarly, the competition model and other forms of connectionism are dependent on the input—as opposed to internal linguistic knowledge—in order to understand how learners create L2 knowledge. On the other hand, a Universal Grammar (UG) approach views input from a different perspective; input is fodder for an internal linguistic system. Schwartz (1993) was clear on this point, at least with regard to syntax; she claimed that new syntactic knowledge grows rather than being learned. Thus, in Schwartz’s view, the input feeds or nurtures an innate system to aid its growth. Primary linguistic data (language exposure) “are necessary for growth of the system of linguistic knowledge” (p. 148). We view input as essential to learning and side with those who argue for a role involving aspects such as frequency, although we are not convinced that that is a sufficient explanation for all aspects of SLA (see Gass & Mackey, 2002).

**Interaction**

In recent years, there has been growing interest in the role of interaction in SLA. Interaction refers to exchanges in which there is some indication that an utterance has not been entirely understood. This can be seen in the interaction between a native speaker (NS) and a nonnative speaker (NNS) from Mackey (1999), given in (1).

(1) NS:  *there’s there’s a pair of reading glasses above the plant*
  
  NNS:  *a what?*
  
  NS:  *glasses reading glasses to see the newspaper?*
  
  NNS:  *glassi?*
  
  NS:  *you wear them to see with, if you can’t see. Reading glasses*
  
  →  NNS:  *ahh ahh glasses to read you say reading glasses*
  
  NS:  *yeah*

In (1), there is recognition of a new lexical item following negotiation of that word, a negotiation that takes place in context. This example illustrates how the learner may have used the conversation as a resource to learn the new phrase “reading glasses” that appeared in the first line. In the penultimate line (marked with an arrow), the NNS acknowledges the fact that the new word “reading glasses” came from the interaction; one can reasonably
assume that the recognition of the meaning came as a consequence of the negotiation. The nonunderstanding can be indicated verbally as in (1) or nonverbally, as in (2), which took place in a restaurant in China.

(2) NNS: Would you like smoking or nonsmoking?
NS: Nonsmoking, please.
NNS: I’m sorry, nonsmoking not available. Would you mind smoking?
NS: Yes.
NNS: (Walks to the smoking section.)
NS: (Does not follow.)

In (2), the nonverbal action of not following is what leads the NNS to understand that something has gone awry. Within the interaction hypothesis (Gass, 1997; Long, 1996), the argument is that in both instances there is some indication to the NNS that something is not being understood, although it is not clear (a) whether NNSs understand the precise error (see Mackey, Gass, & McDonough, 2000) or (b) how to fix the error if they perceive its precise location. The argument goes further to state that such indications of nonunderstanding direct attention to a problem area. Long made this clear in his definition of the interaction hypothesis: “negotiation for meaning, and especially negotiation work that triggers interactional adjustments by the NS or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways” (pp. 451–452, emphasis in original). Thus, through negotiation, a learner’s attentional resources may be oriented to (a) a particular discrepancy between what he or she knows about the L2 and what the L2 really is or (b) an area of the L2 about which the learner has little or no information.

In Gass (1997), a model was presented in which interaction serves to draw attention to an unknown part of language. Learning may take place during the interaction. Negotiation, however, may also be an initial step in learning and serve as a priming device, thereby setting the stage for learning rather than being a forum for actual learning. In this initial capacity, follow-up input is often necessary before learners can begin to internalize language information. In this case, the additional input may serve as a reinforcing device in that it provides the learner with evidence to confirm or disconfirm information about particular parts of the language, as seen in Figure 1.

**Attention**

In our brief description of the input-interaction framework, we mentioned the crucial role of attention. The numerous treatments of this concept go beyond the scope of this study. However, we will discuss some of the recent treatments in the SLA literature. Schmidt (2001) stated that attention “is necessary in order to understand virtually every aspect of second language acquisition” (p. 3). He went further to argue that “SLA is largely driven by what learners pay attention to and notice in target language input” (pp. 3–4). Gass (1988,
1997) and Schmidt (1990, 1992, 1993, 1994, 2001) argued that attention is a key factor that determines whether something is noticed in the input. For Schmidt, noticing is a prerequisite for intake. Gass (1997), on the other hand, used the notion of *apperception*. Apperception, or “the process of understanding by which newly observed qualities of an object are related to past experiences” (p. 17), serves as selective cueing for the very first step of converting input into intake in Gass’s proposal. Apperception “relates to the potentiality of comprehension of input” (p. 4) but does not guarantee that it will result in intake.

Although most scholars recognize the importance of attention, it is the nature of attention itself that is at the root of recent debate. A crucial point of controversy is whether attention relies on capacity limits. Robinson (2003) suggested that recent controversies concern the extent to which attentional resources are limited or unlimited. Models in which attentional resources are unlimited, such as the interference model (see, e.g., Gopher, 1992), argue that increasing the number of stimuli and response alternatives or the similarity between them will sometimes lead to confusion, reducing performance efficiency. This can be caused by competition for the same types of codes during information flow, or by “cross-talk” between similar codes. (Robinson, pp. 644–645)

Thus, interference becomes a matter of involuntary attention shift rather than the result of attentional resource limits.
However, some models view attention as limited in its capacity. As Robinson (2003) noted, the multiple-resource model of attention appeared as a response to the observation that some simultaneous, dual tasks do not affect performance, whereas others do (Wickens, 1989). Breakdowns in performance are dependent on whether the same or different resource pools have been used. For example, attention directed to simultaneous tasks that draw resources from different pools (i.e., audio vs. visual modality) is not likely to have any difficulty or interference. However, when attention is drawn from the same pool, breakdowns in performance are more likely to occur.

In another model, it has been argued that the human brain is equipped with a single, limited attentional processing system that allows for a certain amount of input to be attended to at one time (Anderson, 1983; Kahneman, 1973; Kihlstrom, 1987). Cognitive overload explains why learners who are constantly exposed to a language cannot make use of all available input. That is, learners establish a mechanism that allows them to filter out some language stimuli from the input, which results in input that is ultimately more manageable and easily digestible, thus ready for further processing. What is isolated is clearly dependent on a learner’s stage of development and readiness to learn (Gass, Svetics, & Lemelin, 2003; Mackey, 1999).

Sharwood Smith (1993) referred to input enhancement as a means of highlighting certain language areas for learners with the goal of drawing their attention to those areas. He pointed out that input enhancement can be internally and externally driven. Externally driven enhancement is what happens in a classroom when a teacher, through a variety of means, draws attention to a particular area of language (e.g., through a structured task or an overt explanation). Internally driven enhancement comes about through learners’ own devices when they attend to something themselves (e.g., due to salience or frequency). A question that needs to be addressed concerns the extent to which different parts of language are more or less amenable to internally versus externally driven enhancement.

**Input**

That input plays an essential and necessary role in language acquisition seems to be an unquestionable fact. What is far from being uncontroversial is how input is transformed into intake. Processing instruction (VanPatten, 2002) “is a type of grammar instruction or focus on form derived from the insights of [input processing]” (p. 9). The input processing (IP) model proposed by VanPatten (1990, 1996; see also work on processing instruction first discussed by VanPatten & Cadierno, 1993) focuses on the need to understand learners’ psycholinguistic processing strategies. In particular, this model seeks to explain how learners’ processing strategies may hinder the development of their internal language system, particularly with linguistic data that have certain features—for example, those that lack communicative value. Assuming a limited capac-
ity for attention, this model assumes that—all things being equal—form and meaning compete for attentional resources, with meaning more likely to win out (particularly at early stages of acquisition). According to the IP model, processing strategies may mislead learners in the comprehension of input, which results in the development of impoverished intake and eventually a faulty internal language system. Processing instruction therefore seeks to reorient learners’ attention when processing online input to create meaningful form-meaning connections that eventually lead to adequate intake.

Input processing is not uncontroversial. DeKeyser, Salaberry, Robinson, and Harrington (2002) critically examined the theoretical grounds in VanPatten’s model of input processing and argued that there are problems in the conceptualization of attention and input processing. A single, limited attentional capacity model as proposed in VanPatten’s model does not specify how and why attention is constrained in capacity and leaves unexplained how exactly processing operates on language. DeKeyser et al. also argued that “the status of IP as a psycholinguistically testable construct is questionable” (p. 6), mainly because of the difficulty of assessing how learners process input with a meaning-driven parsing mechanism and because this meaning-driven parsing mechanism, although possible, contradicts current sentence-processing approaches and ignores the importance of structural language information for sentence interpretation.

Finally, it is important to note that input may be analyzed at different levels of sophistication. Gass (1988, 1997) proposed that the greater the depth of analysis, the more likely it is that input will become intake. For instance, learners “may have processed [an] utterance simply for meaning and not noticed and stored the different structure manifest in the input” (Sharwood Smith, 1993, p. 168), which clearly indicates that an analysis at this level in comprehension may serve a lesser purpose for intake than an analysis made at the level of morphology, lexicon, or syntax.

**Language Areas**

Gass et al. (2003) found different effects for focused attention depending on the degree of abstractness, complexity, and L1-L2 difference, which they operationalized in terms of different parts of language (e.g., vocabulary or morphosyntax). They argued that learners require less externally focused attention for vocabulary learning because their internal resources are better able to be tapped for concrete and noncomplex language, such as the vocabulary items in their study. Vocabulary—at least in the present study—is also concrete and noncomplex. Thus, vocabulary appears to be an area where minimal input and interaction, or both, is necessary for learning to take place. This idea finds further support in work by Mackey et al. (2000) on learner perception of interactional feedback. The main focus of their analysis was on morphosyntax, the lexicon, and phonology. The extent to which a learner recognized feedback as
it was intended (i.e., lexical feedback perceived as lexical feedback) depended on the language area in question, with only a weak correspondence for morphosyntactic feedback.

The role of different parts of language is not new. VanPatten (1994) suggested that “perhaps different aspects of language are processed and stored differentially” (p. 31), and Schmidt (1995) similarly proposed that “different aspects [of language] may require more or less of [attention]” (p. 14). Schmidt (2001) further argued that the field of SLA “requires research within SLA itself, focused on different domains of language” (p. 24). The current study makes the important assumption, following Long (1996) and many others, that what happens in interactions involves drawing attention to problem areas of learner speech.

The question arises as to why different areas of language should react differently to different types of information (in the present study, input and interaction). We therefore turn to issues of abstractness and complexity. Both DeKeyser (1995) and Schmidt (2001) raised the issue of abstractness in the context of what can be noticed or (in the case of DeKeyser) what can be memorized. Schmidt assumed that “the objects of attention and noticing are elements of the surface structure of utterances in the input, instances of language, rather than any abstract rules or principles of which such instances may be exemplars” (p. 4). DeKeyser used abstractness to predict the likelihood of memorization, which is least likely when a form depends on the abstract characteristics of its environment rather than on a concrete surface element.

Hulstijn and de Graaff (1994) proposed that “the advantage of explicit instruction is greater in the case of complex rules than in the case of simple rules” (p. 103). Complexity is defined as “the number (and/or the type) of criteria to be applied in order to arrive at the corrective form” (p. 103). Even though we are not dealing with explicit instruction in this study, we believe that instruction has parallels with the kind of data that concerns us (i.e., input and interactional data), given that in both instruction and interaction, attention is an underlying principle of explanation. If a particular area of language is simple (i.e., noncomplex), learners can gather appropriate information on their own. When the language form or rule is more complex, external intervention may save learners time (although Hulstijn & de Graaff do not state that learners are unable to figure out complex phenomena on their own). Thus, interaction may serve a more significant function with greater complexity. In other words, some parts of language may be particularly difficult to figure out without some outside help.

**HYPOTHESES**

In the previous sections, we have argued that input and interaction serve different functions with regard to language learning and that these functions differ with regard to different language areas. The present study is concerned
with the order and presentation of language through input and interaction and its effect on different language areas. Our goal was to investigate the effects of material that has (a) an input focus, (b) an interaction focus, (c) an input focus followed by interaction, and (d) an interaction focus followed by input. This study investigated five hypotheses:

1. Given the differential functions of input and interaction, we predict that performance will differ following treatment involving one or the other and performance will also differ following treatment in which the order of presentation differs.
2. Given that interaction is said to be an attention-drawing device, the three experimental groups with interaction will perform better than the group with no interaction.
3. Because input and interaction serve different yet important functions, when there is a combination of conditions (input followed by interaction or interaction followed by input), performance will be better than when only one type of presentation is available.
4. Given Gass’s (1997) assumption that interaction serves as a priming device that readies learners to utilize follow-up input, the learners who experience interaction followed by input will perform the best in the experimental measures.
5. Because learners behave differently with regard to different language areas, each of the three language areas investigated will yield distinct results. Moreover, hypotheses 1–4 will be applicable for each language area.

METHOD

Participants

The original pool of participants in this study consisted of 110 English-speaking learners of L2 Spanish. They were in their third semester of study at a large Midwestern university. Eight learners reported knowledge of another language and were thus excluded, leaving a final pool of 102 learners, all monolingual speakers of English. There were 32 males and 70 females, ages 17–36 years, with an average of 3.5 years of Spanish instruction.

Target Structures

We selected two grammatical structures that are widely accepted to be learned late by speakers of English: (a) grammatical gender agreement and (b) the use of the copula *estar*. Both gender agreement and the copula *estar* are forms of high frequency in Spanish. Frequency is usually regarded as a factor that may make a form salient to the learner (Bardovi-Harlig, 1987). However, evidence suggests that despite their frequency in the input, learners of Spanish fail to acquire these forms (Fernández-García, 1999; Geeslin, 2003; Rogers, 1987). This may be explained, in part, by the relatively low communicative value of these forms (Harley, 1998; VanPatten, 1996). In the case of gender agreement, the
adjectival suffixes \(-o\) and \(-a\) do not add meaning to a phrase but only maintain the grammatical marking system. Similarly, the choice of copula (\(estar\) or \(ser\)) “is determined by the intended meaning of the adjective or sentence and very often sentences containing these verbs are interpretable via the content lexical items and the context in which the sentence occurs” (VanPatten, p. 113). Importantly, failure to assign the appropriate gender or to use the correct copula is unlikely to cause a breakdown in communication, although, as we discuss subsequently, there are instances where communication breakdowns are possible—for example, when adjectives change meaning with different copulas.

In addition to their low communicative value, both grammatical gender assignment and a dual copula system are absent in English. For the copula system, the problem is made more difficult by the fact that \(ser\) is far more frequent in the input, which leads to an early and long stage of overgeneralization of \(ser\) and a late appearance of \(estar\) in the developmental stages of language learners (VanPatten, 1987). In short, low communicative value and absence in the first language are factors that, in combination, tend to make these forms highly problematic for English learners of Spanish despite high frequency in the input.

**Gender Agreement.** Spanish uses masculine and feminine gender for both animate and inanimate nouns. Generally, masculine nouns end with \(-o\) and feminine nouns end with \(-a\). The determiner and the adjective that accompany the noun take the grammatical gender of the noun and overtly establish agreement, as in (3) and (4).

\[
\begin{array}{ll}
(3) & \text{Masculine} & \text{Feminine} \\
& \text{el niño alto} & \text{la niña alta} \\
& \text{the-MASC boy-MASC tall-MASC} & \text{the-FEM girl-FEM tall-FEM} \\
& \text{“the tall boy”} & \text{“the tall girl”} \\
\end{array}
\]

\[
\begin{array}{ll}
(4) & \text{Masculine} & \text{Feminine} \\
& \text{el libro rojo} & \text{la lámpara mágica} \\
& \text{the-MASC book-MASC red-MASC} & \text{the-FEM lamp-FEM magic-FEM} \\
& \text{“the red book”} & \text{“the magic lamp”} \\
\end{array}
\]

Many nouns, however, do not end with \(-o\) or \(-a\) but rather with a consonant or other vowel, yet agreement must be established with both the determiner and the adjective that modify the noun, as in (5) and (6).

\[
\begin{array}{ll}
(5) & \text{Masculine} & \text{Feminine} \\
& \text{el bigote pequeño} & \text{la llave rota} \\
& \text{the-MASC moustache-MASC small-MASC} & \text{the-FEM key-FEM broken-FEM} \\
& \text{“the small moustache”} & \text{“the broken key”} \\
\end{array}
\]

\[
\begin{array}{ll}
(6) & \text{Masculine} & \text{Feminine} \\
& \text{el árbol viejo} & \text{la luz blanca} \\
& \text{the-MASC tree-MASC old-MASC} & \text{the-FEM light-FEM white-FEM} \\
& \text{“the old tree”} & \text{“the white light”} \\
\end{array}
\]

**Attention When?**
**Estar + Location.** *Ser* and *estar* represent a complex dual system in which the choice of copula may entail a subtle aspectual change in the semantics of an adjective. In sentences such as *soy aburrido* “I am boring” and *estoy aburrido* “I am bored,” the choice of copula determines whether the adjective refers to an inherent quality of the speaker or rather the product of a temporal situation. However, this contrast between *ser* and *estar* is not always straightforward or easy to explain, given the semantic and pragmatic dimensions of copula choice (see Geeslin, 2003). Both native and bilingual speakers show great variation in the selection and use of *ser* and *estar* in adjectival contexts, which suggests an ongoing and steady language-change process (Geeslin, 2001, 2002a, 2002b).

Although *ser* and *estar* have many uses, for this study, we specifically targeted the use of *estar* to indicate location when followed by a prepositional phrase.\(^3\) According to studies of stages of development for copula selection (Gunterman, 1992; Ryan & Lafford, 1992; VanPatten, 1987), *estar* followed by a prepositional phrase indicating location is one of the latest stages of acquisition.\(^4\) This form is illustrated in (7).

\[(7) \text{Las llaves están encima del televisor.} \]
\[
\text{the-FEM keys-FEM are on the-MASC television set-MASC}
\]

“The keys are on the television set.”

**Vocabulary**

We selected vocabulary items for the current study following a pilot study with learners of Spanish in their fourth semester \((n = 71)\) that indicated that a low number of participants (between 4% and 35%) knew the following words: *jarrón* “vase,” *cajón* “drawer,” *bigote* “moustache,” *taza* “cup,” *maleta* “suitcase,” *joyero* “jewelry box,” and *cómoda* “dresser.” Because the learners in this study were of a lower proficiency level, we were confident that these same words were unlikely to be known by most learners in their third semester.

Given the earlier discussion regarding abstractness and complexity, we categorize these parts of language in Table 1.

**Materials**

**Pretest and Posttest.** Two sets of testing materials were used for the pretest and posttest: (a) a pen-and-paper translation for vocabulary and (b) a computer-based acceptability judgment task for the two target language forms (gender agreement and *estar* + location).\(^5\) These are presented in Appendixes A and B (excluding distracters).

The translation portion of both tests consisted of a paragraph in Spanish with the seven target vocabulary items. The maximum score was 7. The structure of both translations (pretest and posttest) was designed to be syntacti-
The other portion of the pretest and posttest consisted of an acceptability judgment test. Participants performed this section of the test on the computer. After listening to instructions and carrying out a practice run with four sentences in Spanish, they entered an identification number that was recorded by the computer. The participants then read a series of 24 sentences in Spanish on the screen and marked “correcto” if they believed the sentence was correct or “incorrecto” if they found an error. Each sentence appeared one at a time, remained on the screen for a maximum of 18 seconds, and automatically proceeded to the next. At the end of the session, a list of sentences that the participant had marked as “incorrecto” was printed, and the participant was asked to correct the sentences. The maximum score was 6 for estar and 12 for gender agreement.

The 24 sentences in each test were drawn from a pool of 48 sentences in the program, all of which were of comparable syllable length and difficulty. The 48 sentences were divided into four equal groups, each targeting one of the following: (a) grammatical gender agreement with nouns ending in -o or -a, (b) grammatical gender agreement with nouns ending in a consonant or vowel other than -o or -a, (c) estar + location, and (d) distracters. For each participant, the computer randomly chose 24 sentences made up of six sentences from each of the four categories. For each identification number, the computer recorded the sentences that a participant received in the pretest, so that during the posttest it could present the learner with the 24 sentences not yet covered. The computer recorded the learners’ responses as well as response time (response time is not addressed in this paper). As such, each participant received an individualized pretest and posttest, thereby minimizing the possibility of a lack of comparability in pretests and posttests.

**Input-Only Group.** The materials were designed as small thematic units that covered two sessions of 20 minutes each. The first session consisted of
three activities and the second of two. Three of the five activities were listening activities, which allowed for identical and consistent input. The target structures and vocabulary were embedded in the input and in the activities in a naturalistic way, and meaning became the central focus of the sessions.

Activities were presented in the form of two instructional packets—one for each session. The central theme was a robbery at a luxury hotel. The goal of the unit was to identify the personal belongings that had been moved or stolen from one of the rooms in the hotel. Learners were introduced to new vocabulary and were exposed to input in which the descriptions and locations of objects were necessary. Descriptions of objects required the use of adjectives that modify nouns (and thus the use of gender agreement), as illustrated in (8), taken from a listening activity.

(8) En la mesa había también un jarrón blanco con un ramo de rosas rojas. “On the table there was also a white vase with a bouquet of red roses.”

Additionally, the description of the location of objects entailed the use of estar, as in (9), from the same listening activity.

(9) Mi joyero estaba encima de la cómoda, justo al lado del teléfono. “My jewelry box was on the dresser, just next to the telephone.”

The first activity was a vocabulary activity. Learners were presented with 14 pictures and their corresponding words, among which were the target vocabulary items; learners had to use their world knowledge to classify objects according to whether these objects are typically stolen (e.g., drawer or dresser vs. cash or vase). In the second activity of the first packet, learners listened to Mrs. Rodriguez’s recorded phone conversation with the police, in which she described where her personal belongings were before the robbery. Learners had the names of the objects (article + noun) and a picture of Mrs. Rodriguez’s room. Once they had listened to the audio text twice, they had to draw the objects in the appropriate location in the picture. The third and final activity involved listening to the same audio text for a third time. Learners had to identify a particular characteristic of the object (i.e., whether the jewelry box was big or small). After listening to the audio text, learners read a set of seven statements and circled the appropriate feature (adjective) that corresponded to the object.

In the first activity of the second session, learners listened to the description of the room after the robbery and repeated the exercise corresponding to the second activity of the previous session (drawing objects in the appro
priate location). The last activity again focused on the vocabulary and compared the before and after pictures of the room. The learners’ goal was to classify objects according to whether they were in the same place after the robbery, were in a different place, or had been stolen.

**Interaction-Only Group.** These materials consisted of communicative tasks that covered two sessions of 20 minutes each. Each session consisted of two communicative tasks (jigsaw and information gap), and two versions of the same task type were employed. In the jigsaw task, two pictures (a living room scene and a bedroom scene) were used. Each participant had a similar picture, and the goal was to work together to discover the differences between the two. There were 10 differences in all. Both pictures required the use of the target vocabulary because all the objects were present in these tasks. There were also at least four instances in each task in which it was necessary to state the color of objects in the two pictures, because some differed by color. For example, whereas one picture had a vase with red flowers, the other had a vase with yellow flowers. Thus, to complete the task, it was necessary to use nouns and adjectives together. Similarly, there were at least four instances in each task in which it was necessary to state the location of some objects, as they appeared in different places in each picture. For example, one picture showed two champagne glasses next to the telephone, whereas the other showed a jewelry box next to the telephone. As such, the task required the use of *estar* + location.

Two pictures were used in the information-gap tasks, in which the goal was for the learner to describe the picture to the researcher so that the latter could draw a picture as similar as possible to the learner’s. One scene depicted a man with a moustache sitting on a chair, holding a cup of tea, and facing a woman who was sitting on the floor with a bottle between her legs and holding a glass. In the other, a man with a moustache was sitting on a couch holding a cup of coffee. A woman was standing behind the couch, holding a bottle in her left hand. In both pictures, the target vocabulary was present. To describe the picture, at least 10 instances required the learner to state the color of objects or describe the characters’ physical appearance. In at least four instances, it was necessary to state the location of objects and people. In sum, to complete the task, it was necessary to use nouns and adjectives together (gender agreement) and to describe the location of the object (*estar* + location).

**Interaction + Input and Input + Interaction Groups.** The treatment materials for the interaction + input group and the input + interaction group were a combination of materials from both the input-only and the interaction-only groups. The input session corresponded with the first session of the input-only group, and the interaction session corresponded with the first session of the interaction-only group. The target structures and vocabulary items were present and balanced in the materials across all sessions.
PROCEDURE

As can be seen in Table 2, the study followed a pretest-posttest design. There were four experimental groups and a control group. All sessions for each group—including the pretest and posttest—were carried out within a period of 1 week. All groups were administered a pretest at the beginning of the week and a posttest at the end of the week. Intact classes were used, randomly assigned into the following conditions:

1. Control group (n = 16): This group was administered only the pretest and posttest.
2. Input only (n = 23): The treatment consisted of two in-class input sessions.
3. Interaction only (n = 26): The treatment consisted of two one-on-one interactional sessions with a researcher.
4. Input + interaction (n = 19): The first session of the treatment consisted of an in-class input session. The second session was a one-on-one interactional session with a researcher.
5. Interaction + input (n = 18): The first session was a one-on-one interactional session with a researcher. The second session of the treatment consisted of an in-class input session.

Pretests and Posttests

A week before the administration of the pretest, those learners who agreed to participate completed a consent form and a background questionnaire. Each learner completed a pretest in which he or she received oral instructions in English on how to proceed with the two activities. Learners in all groups were randomly assigned to begin either with the activity on the computer or with the pen-and-paper activity. The pretest lasted approximately 30 minutes. The posttest was identical in format and length to the pretest.

Treatment Sessions

The same native speaker of Spanish—an experienced instructor and one of the researchers of this study—was responsible for the data collection of all
the sessions across groups. The researcher had no contact with any of these learners prior to the investigation. The treatment sessions for the input-only condition were carried out in the classroom during regular classroom hours. Both sessions were conducted exclusively in Spanish. Instructions were read and clarified as needed by the researcher before the learners engaged in the activities. Additionally, there was no interaction in any of the activities. Once the learners had finished, the researcher went over the answers with them, providing in her responses additional input on the target areas.\textsuperscript{13} When the session was over, learners were required to hand in the instructional packet.

The second session was similar in structure to the first. Learners were given a new packet as well as the one they had completed in the first session, because the last activity of this session required information from the second activity of the first packet. Once all the participants had finished, the researcher went over the answers, again providing additional input on the target areas. For the second and final activity, the researcher guided the group of participants through the activity, again providing additional input.

The treatment sessions for the interaction-only condition were carried out as one-on-one NS-NNS task-based sessions. Prior to the session, the researcher provided the materials (pictures) needed for the task and gave oral instructions in English to the learner. Then the learner and the researcher faced each other at a table and completed the tasks. During the interaction, the researcher provided feedback (i.e., elicitation, recast, or negotiation) on the target forms. These sessions were tape recorded.

Materials and procedures for the interaction + input condition and the input + interaction conditions were the same. The only difference between these two groups was the ordering of the treatment sessions.

RESULTS

The first hypothesis dealt with differences among groups.\textsuperscript{14} To determine the validity of this hypothesis, we subjected the data to an ANOVA, $F(4, 97) = 8.196$, $p = .000$. A post hoc Tukey test showed significant differences only between each of the experimental groups and the control group, $p < .05$. None of the experimental groups differed significantly from one another. Table 3 shows the descriptive statistics for the gain scores of the four experimental groups (see Tables 4–6 for actual pre- and posttest scores by language area) and of the control group, based on gain scores from pretest to posttest. The second hypothesis involved the interaction groups. The hypothesis was that the three experimental groups with interaction would perform better than the group with no interaction. To test this hypothesis, we used a $t$-test to compare the two groups. However, this hypothesis was not supported.

The third hypothesis concerned a combination of conditions: input and interaction. When both input and interaction are involved (input followed by interaction and interaction followed by input), performance will be better than
when only one type of presentation is available. We examined the effect of this combination in three ways. First, groups with a combination (input + interaction and interaction + input) were compared to both groups with only one type of presentation (input only or interaction only). The results of a t-test did not support this hypothesis, \( t(84) = 1.431, p = .078 \), although it approached significance. Second, groups with a combination were compared with the input-only group. Again, this interaction was not significant. Third, groups with a combination were compared to the interaction-only group. No significant differences were found.

In the fourth hypothesis, we proposed that the interaction + input group would perform better than all the other groups. Because the results from the first hypothesis (differences among groups) were not significant, this hypothesis cannot be supported.

Hypothesis five is the most complicated and the one that most directly addresses the concerns of this study. It has as its scope the investigation of each language area separately. Tables 4–6 show the descriptive statistics (gain scores from pretest to posttest) based on each language area separately. In

<table>
<thead>
<tr>
<th>Table 3. Descriptive statistics (gain scores)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>Input + interaction (( n = 19 ))</td>
</tr>
<tr>
<td>Interaction + input (( n = 18 ))</td>
</tr>
<tr>
<td>Input only (( n = 23 ))</td>
</tr>
<tr>
<td>Interaction only (( n = 26 ))</td>
</tr>
<tr>
<td>Control (( n = 16 ))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4. Descriptive statistics for vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>Input + interaction (( n = 19 ))</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Interaction + input (( n = 18 ))</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Input only (( n = 23 ))</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Interaction only (( n = 26 ))</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Control (( n = 16 ))</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Note. Standard deviations appear in parentheses.*
what follows, we will reconsider the first four hypotheses in light of the fifth hypothesis, which predicted a difference among groups.

The first hypothesis anticipated differences among groups. We first examine vocabulary (see Figure 2). Table 7 shows the results of paired t-tests with associated effect sizes. With the exception of the control group, all groups significantly improved from pretest to posttest and have medium to large size effects. An ANOVA revealed significant differences between each experimental group and the control group, \( F(4, 97), p = .000 \) (post hoc Tukey, \( p < .05 \)). All groups have medium to large effect sizes. No significant differences were found between any of the experimental groups. Thus, for vocabulary, the first hypothesis is not supported.

For gender agreement (see Figure 3), paired t-tests (see Table 8) indicated a significant difference from pretest to posttest for the interaction + input group,
$t(17) = 2.28$, $p < .05$, and the input + interaction group, $t(18) = -1.81$, $p < .043$, whereas the other two experimental groups did not show significant improvement.$^{16}$ All effect sizes were small. An ANOVA revealed nonsignificant differences, $F(4, 97) = .907$.

For *estar* + location (see Figure 4), paired $t$-tests (see Table 9) indicated significant learning for the interaction + input group only, $t(17) = 3.17$, $p = .003$. A medium effect size exists for that group but not the others. Interestingly, the interaction-only group showed a decrease from pretest to posttest. An ANOVA yielded a significant $F$ value, $F(4, 97) = 2.666$, $p = .037$. A post hoc Tukey test showed a significant difference between the interaction + input

![](image)

**Figure 2.** Mean vocabulary gain scores.

<table>
<thead>
<tr>
<th>Group</th>
<th>$M$</th>
<th>$SD$</th>
<th>$df$</th>
<th>$t$</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input + interaction</td>
<td>-4.26</td>
<td>1.44</td>
<td>18</td>
<td>-12.84*</td>
<td>0.90</td>
</tr>
<tr>
<td>Interaction + input</td>
<td>-3.72</td>
<td>2.13</td>
<td>17</td>
<td>-7.39*</td>
<td>0.76</td>
</tr>
<tr>
<td>Input only</td>
<td>-3.08</td>
<td>1.83</td>
<td>22</td>
<td>-8.08*</td>
<td>0.74</td>
</tr>
<tr>
<td>Interaction only</td>
<td>-3.38</td>
<td>1.60</td>
<td>25</td>
<td>-10.77*</td>
<td>0.82</td>
</tr>
<tr>
<td>Control</td>
<td>-0.25</td>
<td>1.12</td>
<td>15</td>
<td>-0.889</td>
<td>0.05</td>
</tr>
</tbody>
</table>

* $p < .001$. 

---

Susan M. Gass and María José Alvarez Torres
group and the input + interaction group, with the former outperforming the latter, \( p < .05 \). Additionally, this same test substantiated a significant difference between the interaction + input and the interaction-only groups, again with the former outperforming the latter, \( p < .05 \).

According to the second hypothesis, the three interaction groups (interaction + input, input + interaction, and interaction only) would outperform the input-only group. This was tested using a \( t \)-test. This was not significant in any of the language areas, although the results approached significance with vocabulary, \( t(84) = 1.50, p = .07 \).

| Table 8. Paired one-tailed \( t \)-test for gender gain scores by group (pretest-posttest) |
|---------------------------------|--------|------|------|-----|------|
| Group                          | \( M \) | \( SD \) | \( df \) | \( t \) | Effect size |
| Input + interaction            | -0.94  | 2.27  | 18    | -1.81| 0.15   |
| Interaction + input            | -1.33  | 2.47  | 17    | -2.28*| 0.23   |
| Input only                     | -0.43  | 2.40  | 22    | -0.86| 0      |
| Interaction only               | -0.15  | 2.89  | 25    | -0.27| 0      |
| Control                        | -0.06  | 2.23  | 15    | 0.11 | 0      |

\(*p < .025.\)
According to the third hypothesis, two types of presentation conditions could be more beneficial than one, even when time on task is controlled for. Again, $t$-tests were used. We tested for a combination versus input only and a combination versus interaction only. In the case of vocabulary, differences were significant only for a combination versus input only, $t(58) = 1.840, p = .03$, but approached significance for a combination versus interaction only, $t(61) = 1.334, p = .09$. No significant differences were found for gender agreement or for $estar + location$, although for the latter, the combination versus interaction only approached significance, $t(61) = 1.447, p = .077$.

In addition to statistical measures, an examination of the notable trends in these data is informative; we feel that it is revealing to see which conditions yielded the most evidence of learning. As can be seen in Figure 5, with the exception of $estar$ (a discussion of which follows), the combination groups outperformed the groups exposed to only a single treatment condition. Figure 5 shows the relative rank of each condition for each language area. In other words, in the interaction + input condition, the greatest amount of learning occurs for gender and $estar$ and the second greatest amount for vocabulary. Thus, with the exception of $estar$ for input + interaction, the greatest amount of learning occurred in the combination conditions.

| Table 9. Paired one-tailed $t$-test for $estar + location$ gain scores by group (pretest-posttest) |
| --- | --- | --- | --- | --- | --- |
| Group | $M$ | $SD$ | $df$ | $t$ | Effect size |
| Input + interaction | 0.31 | 1.63 | 18 | 0.84 | 0.03 |
| Interaction + input | -1.27 | 1.70 | 17 | -3.17$^*$ | 0.37 |
| Input only | -0.34 | 1.66 | 22 | -1.00 | 0.04 |
| Interaction only | 0.23 | 1.75 | 25 | 0.67 | 0.01 |
| Control | -0.037 | 1.50 | 15 | -1.00 | 0.06 |

*$p < .025.$
The fourth hypothesis concerns the superior performance of the interaction + input group. One can consider the results of the first hypothesis (in the discussion of the fifth hypothesis) to determine the validity of this hypothesis. For vocabulary, this hypothesis is not supported. All groups improved significantly and there was no difference among groups. This hypothesis is partially supported for gender agreement: The only significant learning effect was seen in the interaction + input group, although there were no differences among groups. In the case of estar + location, this hypothesis is supported. The interaction + input group was the only group to show a significant learning effect. Furthermore, from the first part of the fifth hypothesis, an ANOVA revealed superior performance for the interaction + input group versus the input + interaction group and versus the interaction-only group.

**DISCUSSION**

As noted earlier, many researchers (e.g., Schmidt, 1995, 2001; Schwartz, 1993; VanPatten, 1994) have suggested that, to understand SLA, it is important to consider different areas of language; for Schmidt, this is particularly so when dealing with concepts such as attention. Empirical work by Gass et al. (2003) substantiated this proposal by showing the effects of attention in two conditions (+focused attention and −focused attention) when dealing with different parts of language. The present study provides further support for this proposal and includes the additional dimension of distinct presentation types (input or interaction) and the effect of their ordering.

Given the hypothesis that information will affect individual parts of language differently, it is not surprising that, when language areas were pooled, the original hypotheses were not supported. However, when looked at by language area (as in the fifth hypothesis), some interesting results obtain.

When looking at each area (vocabulary, gender agreement, and estar), the greatest amount of learning takes place with vocabulary. In all experimental
conditions, there were significant gains from pretest to posttest. This is consistent with the results of Gass et al. (2003), who found that focused attention was more useful for syntax than for vocabulary. In their nonfocused attention conditions, learners acquired more vocabulary than morphosyntax or syntax. It was suggested that learners cannot use their own internal sources for learning in areas that are highly complex and abstract. But with vocabulary, because it is noncomplex and nonabstract, it is possible to use internal resources more efficiently than with other areas of language. Gass et al. noted two possible interpretations of this finding:

One possibility is that attention is not necessary for lexical learning; incidental learning becomes most important in this context (see articles in Wesche & Paribakht, 1999). In other words, lexical learning appears to be an area that can be learned on the basis of one’s own internal mechanisms. Another interpretation is that learners do in fact use attentional mechanisms in learning new lexical items, but that they are able to more readily use their own internal devices to generate attention. (p. 527)

This is not unlike Sharwood Smith’s (1993) notion of internally versus externally driven enhancement. The results of the current study suggest that learners’ own devices (i.e., internally driven enhancement) are most useful with vocabulary, which we suggest is due to the fact that there is little analysis involved in learning the new vocabulary words presented to learners.

Gass (1997) suggested that negative evidence (e.g., interaction) allows learners to notice an error. In other words, it draws attention to a problem area and can be considered a form of externally driven enhancement. Thus, interaction is a first step in the process of learning. But, in many instances, this may not be sufficient, and learners have to search the input for confirming or disconfirming evidence. If this is correct, then it follows that interaction followed by input would support this model of learning. On the other hand, the reverse order—input followed by interaction—does not facilitate learning because there is no opportunity, following the noticing of an error, to confirm or disconfirm with further input. The results of this study partially support Gass’s model for gender agreement and fully support her model for estar, where significant learning was noted only for the interaction + input group. Because gender agreement and estar are more complex and abstract (or both) than vocabulary, they are more likely to require external intervention (in this case, interactional feedback) than the rather simple vocabulary of this study. As Sharwood Smith (1993) noted, learners can make use of internally driven enhancement. That is, attention can be directed to parts of the input on the basis of the learner’s own internal mechanisms. Unknown vocabulary in the input is often noticeable without any external intervention. That is, interaction may be less useful for vocabulary learning because learners make use of internally driven enhancement and not interaction. As is confirmed by the results of the present study, vocabulary learning responds well to any type of language information.
As discussed previously, input and interaction serve different, yet important, roles in L2 learning. One serves to draw attention to errors, and the other is a means of data gathering. As it has been shown, the trend is for the combination of treatment types to yield better learning rates than one treatment only (significance in the vocabulary versus input tasks supports this), although in other cases, this was only a trend, given that the results were not statistically significant in a strict sense.

One more factor to consider is the trend that is exhibited by superior performance of the input-only versus the interaction-only group for gender agreement and estar. One could imagine learning taking place with one or the other—that is, with input only or interaction only. However, if each is significant for learning, then some other mechanism would have to take the place of the one not available. A question arises as to what that mechanism might be. In the case of input only, learners would have to figure out a mismatch between what they say and what is present in the input. This, of course, is not an impossible task; learners do this all the time. However, we argue that this task is more difficult when the error is not made more salient through interaction. With interaction only, it is more difficult to imagine how learners would confirm or disconfirm hypotheses in the absence of input (unless interaction itself fulfills this role). With this sort of division of labor, one could expect (although it is not a hypothesis of this study) that input only would be more beneficial than interaction only. There is a trend in this direction for the gender agreement and estar results, but, consistent with the other results of this study, because vocabulary is less demanding, there appear to be numerous ways for learners to obtain relevant information.

There may be another way of thinking about the potential superiority of input only versus interaction only, at least where greater complexity is involved. Input, as opposed to interaction, may be a purer source of information, a source not contaminated by noise (in this case, interaction). A learner engaged in interaction has to do a significant amount of work to determine the target of the interactional feedback, and even then, what he or she believes to be the focus of the feedback may not correspond to the interlocutor’s intended focus. As Mackey et al. (2000) showed, there are many possibilities for the interpretation of feedback. Despite the intent of the feedback, its interpretation is not always obvious. With input, the learner’s task is in some sense simpler. What a learner has to do when listening (or reading) is to notice something in the input; there is no interactional noise. If this interpretation is correct, we need to ask why this was not the case for vocabulary. As discussed previously, we suggest that learning vocabulary appears to differ from learning other parts of language and that this has an effect on the resources needed for vocabulary learning.

The current study suggests that among the factors determining learning are notions of complexity and abstractness. As we discussed earlier, the vocabulary words in the present study were not complex and had counterparts in the native language of the participants. Thus, what one had to learn was the
new phonological form for each word. Gender agreement is more complex than vocabulary—it is a new concept for native speakers of English but is not complex in that there is little semantic information to deal with. The amount of learning was less than with vocabulary. What is most interesting is that it is only with gender agreement that the amount of gain reflects the predictions of the model (interaction + input → input + interaction → input → interaction).

Estar location is both complex and abstract. It is complex in that there are morphological changes in the copula itself and abstract in that semantics come into play. Learners have to select from two copulas by determining which is appropriate for a particular context. The differences are subtle and particularly difficult for English learners of Spanish. These difficulties in complexity and also L1-L2 differences yield results that are different from gender or vocabulary. Focused attention (the interaction + input condition) is the treatment that most aided the improvement in estar; the other conditions suggest little other than arbitrary results (e.g., the control and input-only groups were at similar levels). The difficulties in learning estar and the unpredictable results are corroborated by the results of Cheng (2002), who found that estar responded differently to input processing instruction than what would be expected on the basis of numerous other studies within that framework. As noted previously, the Spanish copula is learned late and, of the language areas examined in this study, it appears to be the least amenable to external information. In fact, the only experimental condition that yielded significant learning was that predicted by Gass’s (1997) model: interaction + input.

CONCLUSION

We have attempted to refine the role of input and interaction by showing that their relationship to one another is not simple nor straightforward. We have shown that the way language information is presented to learners and the resulting learning depend on the nature of the target language form. We have considered this in terms of the complexity and abstractness of the form being learned. The less complex the form, the more one can avail oneself of internally driven enhancing mechanisms. The more complex or abstract, the more other sorts of externally driven enhancement devices are used. In general, we can conclude that interaction, as an attention-drawing device, followed by input, as a forum for data gathering, is the most powerful of the externally driven enhancements.

We acknowledge that this study leaves a number of unanswered questions. First, because we were not able to tightly control each interaction, given that each interaction is dependent on what the learner said, we do not know where the question of quantity fits in. That is, is there a threshold for the tokens involved in each of these forms? Does learning depend on pure quantity and, if so, is there an important dividing point? Another area for future research
concerns feedback types. Does the preponderance of one type of feedback over another (e.g., recast or clarification request) yield greater learning? These are not new questions, but they need to be investigated in terms of actual learning rather than in terms of immediate uptake (e.g., Lyster & Ranta, 1997). Finally, follow-up testing was not possible in this particular study. This will be important in future studies as we attempt to determine the role of and inter-relationship between input and interaction.

(Received 26 April 2004)

NOTES

1. One could argue, as did DeKeyser, Salaberry, Robinson, and Harrington (2002), that “attending to form and content during processing for meaning during communicative interaction is a single task drawing on the verbal encoding resource pool . . . not a dual and that simultaneous attention to form and content is clearly possible” (pp. 808–809).
2. This level was selected on the basis of pilot studies in which it was determined that the materials were too difficult for second-semester learners and too easy for those in their fourth semester.
3. There is an exception to the estar + location rule: When the location is an event, ser is used. Our study contained no such examples.
4. Results in VanPatten’s (1987) study identified estar + locatives as stage 4. However, results in Ryan and Lafford (1992) are categorized as stage 5 (the last stage). According to Ryan and Lafford, differences in results may be due to the differences in the nature of the research environment (foreign language instruction vs. study abroad).
5. An anonymous SSLA reviewer questioned the use of two different testing methods. For vocabulary, an acceptability judgment test could not be used because we are not dealing with acceptability or grammaticality, but word meaning. The question remains as to why a translation test was not used for the estar and gender constructions. This was not a possibility, given the difficulty in eliciting appropriate structures, if, following Mackey and Gass (2005), elicitation does not force the structure. For example, if the sentence to be translated had been “The man is in his house” and the learner had written El hombre abita en su casa “The man lives in his house,” we would have no basis for determining what this learner did or did not know about estar (see Gass, 1994, for further discussion of the need for forced elicitation tasks and, particularly, acceptability judgments).
6. This time limit was established in a pilot test.
7. Treatment materials are available from the authors on request.
8. The idea for this unit was adapted from Riggenbach and Samuda (1997, p. 327).
9. Participants in this study did not have class on Wednesdays; thus, for those students who had their posttest at the end of the week, there was no additional classroom exposure for any of the targeted areas.
10. The study was initiated in the 6th week of the semester and took 3 weeks to complete, although each experimental group completed all tasks and tests within a 1-week timeframe. This was done for logistical reasons.
11. During the time of data collection, none of the classes received instruction on either gender agreement or the appropriate use of estar + location.
12. We are aware that input and interaction are potentially confounded with group and individual presentation. Ideally, this could be controlled for, but given the practical restrictions on conducting this type of research, we could not. Because there was no feedback in the input condition, it seems unlikely to us that giving a reading passage or a listening passage in a group or in an individual environment would yield an important difference. Further, we do not believe that this made a difference because the results are not consistent for each of the target areas (vocabulary, gender, and estar + location). For example, for vocabulary, the interaction group did better than the input-only group, but the opposite order was found for the other groups. This could mean that there is an interaction between individual versus group and language area. This is an area for further investigation.
13. The guiding principle in developing the materials was to make them as natural as possible given that they were being used as a classroom exercise. As a result, there was not, for example, absolute equivalence of instances of each vocabulary item; the results are presented in terms of overall vocabulary results rather than results for each word. In other words, we used an input flood model (Trahey & White, 1993), rather than making sure that there was an exact equivalence of all tokens.

14. In the case of t-tests, one-tailed tests were used given the nature of the hypothesis. For the t-value for independent samples in all instances, Levene’s test for equality of variances showed equal variances for all samples. In the case of paired t-tests, Brown (2001) stated that “if the sample sizes are the same for the two groups involved, the t-test is robust to violations of the assumption of equal variances” (p. 156).

15. Because our research focused on improvement resulting from our treatment, our analyses are conducted by means of t-tests and ANOVAs. Effect sizes are presented to indicate the magnitude of the effect and for comparison purposes. The effect sizes were calculated using Cohen’s d-index. We use the standard interpretation of effect size of greater than .8 to be a large effect size, greater than .5 to be a medium effect size, and greater than .2 to be a small effect size.

16. An anonymous SSLA reviewer suggested that we separate out the two noun types (those ending in -a or -o and those ending in a consonant or other vowel). As we have already noted, our words were selected from a larger list of words that were given to a more advanced group of learners; from that list, we selected the seven words that were the least known. This left us with four of the -o/a type and three of the consonant or other vowel type. Given the small number of tokens, we do not feel comfortable making generalizations. What is interesting is that, in three of the four experimental conditions, there was greater improvement with the consonant or other vowel nouns (irregular) than with the -o/a nouns (regular). Given the small sample and the fact that this study was not designed around this difference, this can only be taken as suggestive but is worthy of further investigation.

17. This was not part of any original hypothesis and so was not originally tested for significance. A post hoc analysis revealed no significant differences.

REFERENCES


**APPENDIX A**

**READ THE FOLLOWING TEXT AND TRANSLATE INTO ENGLISH**

En la fiesta de cumpleaños de mi hermana, sus amigas escondieron los regalos en su habitación. Había un regalo que estaba dentro de un jarrón. Era un anillo muy bonito. También había un regalo en un cajón de la cómoda. Era un álbum. Además, la maleta que estaba debajo de la cama también escondía un regalo sorpresa. Era un ratón nuevo para su computadora. Mi hermana también tiene un joyero y allí había otro regalo. Esta vez era un CD de su cantante favorito. El último regalo fue muy simpático. Era un bigote y una nariz de payaso para una fiesta de disfraces. Escondieron la tarjeta de felicitación mientras buscaba los regalos. La tarjeta estaba debajo de una taza en la mesa de trabajo. ¡Fue una fiesta muy divertida!

**TRANSLATION**

At my sister’s birthday party, her friends hid the gifts in her bedroom. There was a gift inside a vase. It was a very nice ring. There was a gift inside a dresser drawer as well. It was a photo album. In addition, the suitcase under her bed hid another surprise gift. This was a new mouse for her computer. My sister also has a jewelry box and there was another gift there. It was her favorite singer’s CD. The last gift was a very funny one. It was a moustache and a clown’s nose for a costume party. They hid the birthday card while she was looking for the gifts. The birthday card was under a cup on her desk. It was a very fun party!

**READ THE FOLLOWING TEXT AND TRANSLATE INTO ENGLISH**

En los grandes almacenes de la ciudad, mis padres hicieron algunas compras con mi novia. Compraron un jarrón que era de varios colores. Era un artículo muy exótico. También compraron una cómoda con varios cajones a un buen precio. Buscaron las ofertas. Por ejemplo, unas tazas que estaban de rebajas también fueron una compra excelente. Compraron una maleta nueva de color negro. Mi madre también compró un joyero y además compró una alfombra. Estas compras eran una sorpresa para nuestra
casa nueva. El último regalo fue muy emocionante. Era un marco con muchas fotografías de mi familia para la estantería del salón. Hay una fotografía de mi padre cuando tenía 20 años. Mi padre tenía un bigote enorme. ¡Es una fotografía muy simpática!

**TRANSLATION**

At the town mall, my parents did some shopping with my girlfriend. They bought a vase that had various colors. It was a very exotic item. They also bought a dresser with various drawers at a very good price. They looked for bargains. For example, some cups that were on sale were also an excellent purchase. They bought a new black suitcase. My mother also bought a jewelry box and even a rug. These purchases were a surprise for our new home. The last gift was very exciting. It was a frame with a lot of family pictures for the bookshelf in the living room. There was a picture of my father when he was 20 years old. My father had a huge moustache. It was a very cute picture!

**APPENDIX B**

**SET 1: ESTAR + LOCATION**

1. *El libro de español está encima de la mesa de la oficina.*  
   “The Spanish book is on the office desk.”
2. *Los boletos son al lado de tu pasaporte y cheques de viaje.*  
   “The tickets are next to your passport and traveler’s checks.”
3. *La cafetería es enfrente de tu restaurante favorito.*  
   “The coffee shop is in front of your favorite restaurant.”
4. *Mis compañeros están en la fiesta de graduación de Pedro.*  
   “My classmates are at Pedro’s graduation party.”
5. *Carlos está en el hospital porque ayer tuvo un accidente grave.*  
   “Carlos is in the hospital because yesterday he had a serious accident.”
6. *Cuando llego a casa, mis hijos son en la escuela todavía.*  
   “When I arrive at home, my kids are still at school.”
7. *El batido de leche está dentro del frigorífico en la cocina.*  
   “The milkshake is in the fridge in the kitchen.”
8. *Las raquetas son cerca de tu bolsa y la bolsa de Ramón.*  
   “The rackets are near your bag and Ramon’s bag.”
   “The museum is near the university library.”
10. *Sus padres están en el concierto de música de Pavarotti.*  
    “Her parents are at Pavarotti’s music concert.”
11. *Emilio está en el trabajo aunque hoy es un día feriado.*  
    “Emilio is at work, although today is a holiday.”
12. *Mientras preparo la cena los invitados son en el salón.*  
    “While I prepare dinner, the guests are in the living room.”
SET 2: GENDER AGREEMENT WITH -O/-A

1. *El año próxima iré de vacaciones a México en un crucero.
Next year, I will go on vacation to Mexico on a cruise.*

2. La lectura favorita de Ana es la historia y la cultura Latina.
Anna’s favorite reading is Latin history and culture.

3. *En mi apartamento hay una cocina pequeño y dos salones.
In my apartment, there is a small kitchen and two living rooms.

4. Ayer mi amiga compró un vestido de verano ligero.
Yesterday my friend bought a light summer dress.

5. A mi instructor le gusta mucho la comida mexicana.
My instructor likes Mexican food a lot.

6. *No quiero ir a la fiesta de Elena con el abrigo vieja.
I do not want to go to Elena’s party with the old coat.

7. El desarrollo económico irá a mejor en Perú este mes.
The economic development will go better in Peru this month.

8. La caja mágica del cuento tiene un color y una forma extraña.
The magic box in the fairy tale has a strange color and shape.

9. *La comisaría tiene una oficina amplio con tres teléfonos.
The police office has a spacious office with three telephones.

10. Mañana el jefe visitará el departamento nuevo de ventas.
Tomorrow the boss will visit the sales department.

11. A mi hermano le interesa mucho la literatura americana.
My brother is interested in American literature.

12. *No pude encontrar a la amiga de Pedro en el museo.
I could not find Pedro’s friend in the modern art museum.

SET 3: GENDER AGREEMENT WITH CONSONANT AND VOWELS OTHER THAN -O/-A

1. El salón del hotel está reservado para nuestra fiesta.
The hotel room is reserved for our party.

2. *La calle oscuro tiene un aspecto triste y desolador.
The dark street has a sad and distressed look.

Susan M. Gass and María José Alvarez Torres
3. Hace dos semanas pasamos una noche aburrida en el teatro.
   “Two weeks ago we spent a boring night at the theatre.”
4. *El traje tiene un cinturón negro para el pantalón de lana.
   “The suit has a black belt for the wool pants.”
5. *En las playas de Cancún pasamos unas vacaciones divertidos.
   “On the beaches of Cancún we had a fun vacation.”
   “Hockey and football can be aggressive sports.”
7. El tacón alto del zapato está hecho con varios materiales.
   “The shoe’s high heel is made with various materials.”
8. *La llave pequeño abre la puerta trasera y delantera.
   “The small key opens the front and back door.”
9. Durante los inviernos miramos la nieve fría desde la ventana.
   “During the winters we look at the cold snow from the window.”
10. *La empresa compra mucho algodón blanca para la confección de ropa.
    “The company buys lots of white cotton for manufacturing clothes.”
11. *En un debate sobre política necesitamos muchas razones válidas.
    “In a debate about politics we need lots of valid reasons.”
12. Mi madre y yo queremos hacer algunos viajes fantásticos.
    “My mother and I want to take some fantastic trips.”