THE ACQUISITION OF SPANISH COPULA CHOICE AND ITS RELATIONSHIP TO LANGUAGE CHANGE

Kimberly L. Geeslin

Indiana University

This paper argues that crossing the boundaries between sociolinguistics and second language acquisition (SLA) is essential to both fields of study (Giacalone Ramat, 1995; Preston, 1993). Specifically, data collected in an investigation of the SLA of copula choice by 77 English-speaking learners of Spanish are examined in terms of similarities to data collected in studies of language change (Gutiérrez, 1992; Silva-Corvalán, 1994). The variables used to analyze these data, Frame of Reference and Susceptibility to Change, have been shown to be useful for examining the process of the extension of *estar* to new [copula + adjective] contexts from a sociolinguistic perspective. The application of these variables to SLA data allows an investigation of the mirror-image relationship predicted to exist between the process of language loss and the process of language acquisition.

Recent advances in sociolinguistics and in SLA have shown that the two fields share important assumptions and often reach conclusions that are mutually relevant (Giacalone Ramat, 1995; Preston, 1993, 1996). In particular, both areas of study refer to universal processes to explain the data and results that are obtained from individual analyses. These universal processes explain speak-

419

ers' tendencies toward certain forms, the early loss of some structures, and the difficulty of acquiring others. Furthermore, sociolinguists and SLA researchers both study language change in terms of movement toward or away from a particular target form. It has even been predicted that a mirror-image relationship exists between the process of language acquisition and language loss. Given the mutually relevant assumptions shared by the fields, the current investigation presents an analysis of the acquisition of copula contrast in Spanish in light of recent sociolinguistic research. Copula contrast in the [copula + adjective] context has been found to be undergoing a process of language change in both bilingual and monolingual populations. Sociolinguists have not only selected relevant variables to describe this change, but they have also succeeded in describing stages of change or loss in terms of such variables (Gutiérrez, 1992; Silva-Corvalán, 1994). The current research evaluates data from second language learners of Spanish using sociolinguistic variables in order to compare the paths of language change with those of language acquisition. This analysis makes it possible to evaluate whether the hypothesized mirror-image correspondence between sociolinguistics and SLA is an accurate description of the relationship between these two processes.

BACKGROUND

A Case for Cross-Disciplinary Investigation

Despite an overall lack of studies that cross boundaries between sociolinguistics and SLA, several researchers have argued for the use of findings in one field of linguistics to improve study in a related area. Romaine (1988, 1994) supported the use of data from pidgin and creole studies to inform theories of language change because the recycling of so-called cultural grammar in languages with a long written history, such as English, is likely to obscure the universal processes involved in change. Therefore, she claimed that, to investigate the processes that underlie change, comparison across many developmental continua is essential (Romaine, 1988, p. 65). One example of the application of this approach is the work of Andersen (1979, 1981). In examining data from SLA (Schumann, 1978) and from pidgins (Bickerton & Odo, 1976), Andersen stated that researchers in the two subfields are actually interested in the same phenomenon but from different points of view. According to him, nativization, or acquisition toward an internal norm (i.e., a pidgin), and denativization, or acquisition toward a target norm (i.e., late SLA), are processes that coexist (see Romaine, 1988, for discussion). Similarly, Greenberg (1991) claimed that the difference between SLA and historical linguistics is that the former is microdiachronic and the latter is macrodiachronic. Thus, the processes that bring about change in each are related. He pointed out that universal principles do not apply in a historical void; one can therefore investigate such principles without requiring that they be exhibited in exactly the

same way. Silva-Corvalán (1991) added further impetus to this argument, stating that processes such as simplification are characteristic of a number of language processes, including first and second language acquisition, foreigner talk, pidginization, creolization, and language loss (p. 325). Concurrently, Williams (1989) compared pidgins, creoles, second languages, nonnative institutionalized varieties of English, dying languages, and the L1 of immigrant communities along sociolinguistic, linguistic, and acquisitional dimensions showing many areas of crossover (see Hawkins, 1991, for a discussion of typology and SLA; see Mufwene, 1990, and Odlin, 1992, for a discussion of SLA and creoles with a special focus on transfer). The fact that many of the stages in the various processes above seem to mirror one another supports the usefulness of exploring this type of relationship.

Romaine (1988) stated that theories of historical linguistics assume that change is regular and governed by constraints, that processes operating today are like those that operated throughout history (uniformitarianism), and that languages are continuously facing forces of internal change. It is necessary to accept these claims for both SLA and natural language change in order for connections to hold between the two. However, the connection to SLA does not necessitate that L2 learners have access to Universal Grammar (UG) because the universal processes may well be constrained by innate knowledge other than UG (such as the subset principle and the noun phrase acquisition hierarchy) as well as by human processing capabilities. In fact, Preston (1996) mentioned that the concerns of SLA researchers should be wider than UG issues because it is improbable that all aspects of learner language fall out from UG, even if adult learners do have access to it. Furthermore, the concept of universal processes does not exclude the fact that change in a naturalistic setting is influenced by external factors (Milroy, 1992).

The Relationship between SLA and Language Change

In contrast to the growing support of interdisciplinary investigation in linguistics, the relationship between SLA and language change has received less attention. Preston (1989) was the first widely known attempt to map out areas where sociolinguistics and SLA could inform each other. The fact that variation in learner speech can be vertical, related to levels of development, or horizontal, related to situational factors, has sometimes impaired communication between SLA researchers and sociolinguists (see Adamson & Regan, 1991, for a discussion of vertical and horizontal variation). In response to this potential confusion, Preston (1993) pointed out that the misunderstandings that exist regarding concepts, findings, and research tools constitute one of the largest impediments to communication between the two groups. Preston (1996) affirmed that variation is affected by (a) the linguistic context, (b) the stylistic level of the discourse, (c) the social identity of the speaker, and (d) the historical position of the speaker and the discourse. Nevertheless, Preston (1993) supported the investigation of vertical variation by describing L2 learn-

ers as being on a "fast track" of language change. He added that L2 learners'
speech may exhibit patterns in a way that is not affected by other influences
as it might be in an L1 speech community because L2 learners are less likely
to attach socially symbolic meanings to variables (1993, p. 164). Berdan (1996)
reconciled many of these differences by stating that the only difference be-
tween horizontal and vertical variation is whether time is an influencing fac-
tor. In this light, looking at how learners vary across time, or at the stages of
acquisition, does not necessitate overlooking other contextual features that
may contribute to social variation. Furthermore, addressing factors in the ex-
ternal environment should not prevent the investigation of the language sys-
tem itself.

   One study that exemplifies the potential for the interdisciplinary study of
language change and SLA is Giacalone Ramat (1995). In her article about gram-
maticalization, Giacalone Ramat stated that language change is connected to
stages of acquisition because the general forces acting on each cause similari-
ties (p. 119). Specifically, she claimed that the similarities between the two are
a result of principles of linguistic encoding and universal cognitive abilities.
She related the process of grammaticalization in historical linguistics, in which
an item loses lexical properties in favor of syntactic ones, to SLA. According
to her, grammaticalization in SLA is the continuous process of the acquisition
of constraints and the shift away from the initial lexical categorization that
learners give to novel items. For example, Giacalone Ramat showed that in the
acquisition of the tense-aspect category by L2 learners of Italian, learners be-
gin to express tense and aspect solely with lexical items and not by morpho-
logical markings. Subsequently, some morphologically marked auxiliary verbs
appear. In the development of the imperfect for example, some learners first
build periphrastic forms with the imperfect of an auxiliary before developing
bound morphology that marks the main verb directly. This process of devel-
opment corresponds with the historical development of auxiliaries in the Ro-
mance and Germanic languages.[1]

   In research directly related to Spanish, Silva-Corvalán (1991) discussed the
process of simplification found in many languages, as exhibited in the tense-
mood-aspect (TMA) system of second- and third-generation bilinguals in the
United States. Her findings were based on interviews with Mexican-Americans
living in Los Angeles who represented a variety of ages and levels of educa-
tional attainment. She found support for a shift to a more isomorphic system,
such that the most frequent meaning of a form became the only use of that
item (see also Dorian, 1980).[2] In a related study, Myhill (1991), using a textual
analysis of various creoles to identify common features in each, showed that
the TMA features believed to characterize creoles were also found in other
languages and in SLA data. Some of the aspects investigated were the marking
of anteriority, future or irrealis, and nonpunctuality. Although it is unclear if
the relationship between creoles and SLA exists because creolization involves
SLA or because creoles and SLA involve the same processes, the similarities
remain valuable. The aforementioned studies show both the potential of theo-

ries that relate language change and SLA and the need to test such theories with additional data. Given this necessity, the following hypothesis was used to guide the current investigation: The process of language change may mirror the process of language acquisition because the former is moving toward simplification and the latter is moving away from simplification to a more native-like grammar.[3]

The Copula Contrast in Spanish

To further specify the above hypothesis, it is necessary to briefly outline the difference between the two Spanish copulas *ser* and *estar*, which represent two distinct meanings in Spanish. Although in some cases a grammatical context allows only one copula, there are other contexts, such as the copula + adjective structure, that allow both. It is these contexts that are most difficult to acquire and are also undergoing a process of language change in monolingual and bilingual populations. With adjectivals, *ser* expresses a quality of the subject; in contrast, *estar* with adjectivals describes a state of being (Solé & Solé, 1977, p. 250). In short, when *ser* appears with adjectives it generally expresses inherent or permanent characteristics, whereas *estar* expresses a physical or mental state of being. In cases where both copulas are permissible, there is a change of meaning depending on which copula is chosen. This is illustrated by examples such as *es borracha* "she is a drunk" versus *está borracha* "she is drunk," and *es alegre* "he is a happy person" versus *está alegre* "he is happy (right now)".

Copula Contrast in Spanish: A Change in Progress

In many varieties of Spanish, *estar* appears with adjectives that were formerly restricted to use with *ser* only. For example, a sentence like *el niño está alto* "the child is tall" is currently used to express the meaning "the child has become tall" even though the adjective *alto* "tall" more often appears with *ser* to express an inherent or permanent characteristic. Sociolinguists have examined this innovative use of *estar* in bilingual populations in Los Angeles (Silva-Corvalán, 1986) and in monolingual populations in Mexico (De Jonge, 1993; Gutiérrez, 1992) and in Caracas (De Jonge).

Silva-Corvalán (1986) examined the extension of the usage of *estar* in the bilingual community of Los Angeles with the goal of representing linguistic change as a series of stages. She collected data from interviews with bilinguals of Mexican descent representing three generations of arrival to the United States, many age levels, and different ages of exposure to English. Unlike previous descriptions of copula contrast, Silva-Corvalán claimed that traditional binary distinctions such as [± inherent characteristic] could not explain her participants' choice of copula in preadjectival contexts. Instead, citing Falk (1979), she stated that each of the contexts in which the copulas *ser* and *estar* appear should be seen as composed of a bundle of features, any of which may

Table 1. Stages of change in the copula + adjective structure described by Silva-Corvalán (1986)

Features constraining use of *estar*	Ungrammatical example	Permissibility		
		Stage 1	Stage 2	Stage 3
Individual frame of reference	**Juan está alto* "Juan is tall" (comparison to group norm)	No	Yes	Yes
Susceptible to change	**Juan está inteligente* "Juan is intelligent" (inherent characteristic, not changing)	No	No	Yes
Circumstantial attribute	**Juan está simpático* "Juan is nice" (defining characteristic, not circumstantial vs. *calvo* "bald")	No	No	Yes

gain prominence according to the interpretation the speaker chooses to highlight. Thus, for monolinguals, the choice between *ser* and *estar* should indicate a subtle semantic difference. This difference is expressed through the association of particular discourse features with a specific copula. Prior to analysis of the data, Silva-Corvalán stated that her expectation was that the contextual features "class" versus "individual" frame of reference, and "susceptible" versus "nonsusceptible" to change, would best explain copula choice (p. 95). In a class frame of reference, a subject is referred to as a member of a group with that attribute (e.g., *es* [*ser*] *alegre* "he is a member of a class of happy people") whereas an individual frame of reference compares a subject to itself at another point in time (e.g., *está* [*estar*] *alegre* "he is happy today"). Silva-Corvalán found that three contextual features best describe copula choice: (a) whether the adjective is one of circumstantiality, (b) whether the adjective is susceptible to change, and (c) whether the speaker chooses to impose a class or an individual frame of reference. Using these features, she characterized the changing use of *estar* as a loss of some or all of the restrictions on its use, mentioned previously, and showed the process of language change as a series of stages. In the first stage, *estar* is constrained by pragmatic features such as frame of reference, and by the semantic relationship between the referent and the attribute such as susceptibility to change and circumstantiality of an attribute. In the next stage of change the restriction on the frame of reference is lost, and susceptibility to change and circumstantiality may or may not be relevant. In the final stage of change, *estar* is used to introduce any attribute and all contrast is lost. Assignment of an individual's usage to a stage is not absolute, but rather reflects frequency of occurrence within a given constraint. Table 1 provides an illustration of these stages and each constraint that governs

them, along with examples that are ungrammatical unless each constraint is violated.

In addition to the features used in Silva-Corvalán (1986, 1994) to describe the stages of change, other linguistic features were examined. Each adjective was categorized according to its class (e.g., age, size, color, physical appearance), and it was found that the categories age, description, physical appearance, evaluation, size, color, and perception favored innovative use of *estar*. Silva-Corvalán also examined the animacy of the referent in each copula + adjective context but found no effect for this variable. Finally, the degree of semantic transparency represented by each copula + adjective token was examined, a variable that refers to the degree to which the semantic contrast created by combining a particular adjective with each of the two copulas is easily perceived. The hypothesis regarding this variable is that the greater the degree of transparency, the less likely change is to occur. Silva-Corvalán examined adjectives that showed a shift in modality when paired with each copula, those that change meanings completely, and those that show near synonymy. She found a 72% chance of innovation when an adjective showed some variation in modality, but its meaning did not change completely. Thus, not only are the features used to describe the stages of language change important, adjective class and semantic transparency add further information.

Once findings for bilingual Spanish speakers were made available, sociolinguists conducted similar research with monolingual Spanish speakers to ascertain whether the change reported by Silva-Corvalán (1986, 1994) was merely the effect of contact with English. Gutiérrez (1992, 1994b) reported important research on the extension of *estar* conducted in a monolingual community in Michoacán, Mexico. By using the distinction of group versus individual frame of reference, and by dividing adjectives into various classes, his study was able to investigate the extension of *estar* in similar depth to Silva-Corvalán (1986). Gutiérrez used data from conversations that called for comparisons and perceptions of situations of which the researcher had prior knowledge, making the individual frame of reference possible, and a fill-in questionnaire that allowed for a closer examination of certain contrasts. His findings were very similar to those in other studies. The classes of adjectives that most encouraged innovation were age, size, physical appearance, and evaluation. The process of increased innovation with *estar* in this monolingual population was also explained by an initial weakening of the frame of reference constraint followed by the extension of *estar* in contexts of subjectivity. For example, adjectives like *calvo* "bald" that are not reversible, and therefore cannot be explained by susceptibility to change, were often introduced by *estar* because of their subjective nature (see also Silva-Corvalán, 1994, p. 98). Next, Gutiérrez (1994a) compared the first-generation speakers in Silva-Corvalán (1986) to his work, controlling for social variables so that contact with English presented the only difference between groups, and showed that there was no difference in usage. He concluded that the change was initiated in Mexico, brought to the United States and then accelerated. It was found that

groups with high levels of education and those belonging to the upper socio-economic class were least likely to participate in the process of the extension of *estar*.

These results have been supported by research on the same topic in other research contexts. De Jonge (1993) provided evidence that *estar* is becoming accepted in new contexts with expressions of age through an analysis of samples of educated monolingual speech in Mexico City and in Caracas, Venezuela. He found a higher frequency of innovation in Caracas, but the same process was found to occur in both cities.[4] Sanz and González (1995) investigated the opposition between the equivalents of *ser* and *estar* in Tortosí Catalán. Using data from spontaneous oral production, controlled production, and a grammaticality judgment task from nine speakers drawn from a wide range of age groups, Sanz and González found a process of change similar to that found by Silva-Corvalán (1986, 1994). Interestingly, they found that once the notions of susceptibility to change and circumstantiality associated with *estar* were lost, speakers began to use movement verbs such as *acaminar* "to walk" as in *acamina coixa* "she walks with a limp" to express the same meaning. In other words, once the contrast between the two copulas no longer expressed a difference in modality, new verbs began a process of grammaticalization and were used to express what *estar* once had. Each of these studies affirms that the use of *estar* is currently being extended to innovative contexts and provides examples of linguistic features that have proved useful in describing this change in progress.

The SLA of Copula Contrast in Spanish

The body of research on the SLA of copula choice that exists today is largely focused on the order in which all contexts in which a copula may appear are acquired. These contexts include not only adjectival contexts but also locative and equative constructions. In a study based on data from picture-based story-telling activities and conversations, VanPatten (1985, 1987) studied six learners throughout the course of one academic year. These data allowed VanPatten to arrive at three basic stages, which were expanded to five in VanPatten (1987) with the addition of data from an unpublished study by a graduate student and classroom observations. The stages VanPatten found consist of: omission of any copula (stage 1); *ser* chosen to perform most functions (stage 2); *estar* appearing with the progressive (stage 3); *estar* appearing with locatives (stage 4); and *estar* appearing with adjectives of condition (stage 5).

To extend these claims to other learner populations, Ryan and Lafford (1992) conducted three interviews over the course of 4½ months with 16 students in a study-abroad setting. The stages found were very similar to Van-Patten's (1985, 1987), with the exception of the order in which *estar* was acquired with locatives and adjectives of condition.[5] The universality of these stages of acquisition has been further confirmed for a number of learner popu-

Table 2. Constraints on usage of *estar* based on a
regression analysis from Geeslin (2000)

Constraint	Level			
	1	2	3	4
Grammatical accuracy	X	X	X	X
Adjective class	X	X	X	
Semantic transparency			X	X
Animacy				X
Susceptibility to change	X			
Dependence on experience	X	X		X
Directionality	X		X	
Dynamicity	X	X	X	X
Perfectivity	X	X	X	
Frame of reference			X	X
Task	X	X	X	X
Telicity			X	

lations in a variety of settings (Briscoe, 1995; Gunterman, 1992; Ramírez-Gelpi, 1995).

In an effort to address the fact that traditional binary oppositions, such as [±adjective of condition], fail to capture all of the subtleties of copula contrast, I hypothesized in Geeslin (2000) that many of the social and theoretical linguistic features commonly employed to describe monolingual usage could provide new tools for the analysis of SLA data. Features included in that investigation were taken from work in sociolinguistics (Gutiérrez, 1992; Silva-Corvalán, 1994), studies on aspect (Lema, 1995; Luján, 1981; Schmitt, 1992; Smith, 1997), and theoretical work on Spanish (Clements, 1988; Falk, 1979). Data from 77 English-speaking high school learners of Spanish were analyzed using a regression analysis based on semistructured interviews, a picture-description task, and a contextualized questionnaire.[6] Results indicated that progress across time in SLA, in this case progress across four levels of enrollment, can be described in terms of those features that significantly predict the use of *estar* at each level of development because such features vary from one level to the next. Table 2 presents a summary of the results, indicating each level at which a linguistic feature was a significant predictor of copula choice.

Although Geeslin (2000) found that many of the contextual features under investigation were significant predictors of the usage of *estar*, a number of questions remain unanswered. For example, it is unclear why certain features (e.g., animacy) are significant only at one level of development whereas other features (e.g., dependence on experience) are significant at all but one level of development. These and other variables fail to show consistent movement toward the inclusion or exclusion of a constraint as proficiency increases. Additionally, some features were found to be significant at all levels of development and, although these may represent universal features, they do not

necessarily contribute specifically to the description of change across time. I hypothesized that those features that are only significant at one or two levels of development may indeed interact with other features included in the equation such that both are not necessary to describe learner behavior at that level. In that study I concluded that, although the results point to the importance of an approach that acknowledges the interaction of contextual features to determine copula choice, further investigation of each individual feature is essential to better understand its role in the process of SLA. The current investigation aims to address this problem by using the hypothesis set forth in studies on language change to guide the inquiry. In short, an examination of the predicted mirror-image relationship between language change and language acquisition specifically calls for the examination of those features, and only those features, included in sociolinguistic studies of language change.

THE PRESENT STUDY

Returning to the hypothesis derived from general research on the potential connection between SLA and language change, and taking into consideration the previous research on the SLA of copula choice, it is possible to further specify the research question investigated in the current study. The former hypothesis—that the process of language change may mirror the process of language acquisition because the former is moving toward simplification and the latter is moving away from simplification to a more nativelike grammar—can be rewritten as: The stages set forth for language change may mirror the stages of SLA exhibited in the present study such that the constraints on susceptibility to change and circumstantiality will be acquired prior to the frame of reference constraint.

Participants

The participants in the current study were selected because of their limited contact with native Spanish and the homogeneity of their backgrounds. The limited contact with Spanish was necessary to ensure that the results in the present study were not the product of the acquisition of the sociolinguistic norms of native-speaking populations. Each of the 77 participants was a high school student, recruited for participation by his or her high school Spanish teacher in exchange for extra credit. All students in the course sections involved in the study were invited to participate. Each student had successfully completed all previous levels of instruction.

Participants ranged in age from 14 to 18 years and were enrolled in one of four levels of instruction. Participants varied in terms of the length of time they had studied Spanish from 3 to 6 years. For the purpose of this study, students with 3 years of study are called level 1. It should be pointed out that because 2 full years of instruction are offered prior to enrollment in the high school, the third year of study is comparable in terms of curriculum to a sec-

ond-semester college course. A total number of 77 students volunteered, 20 from level 1, 23 from level 2, 24 from level 3, and 10 from level 4. Low enrollment, and consequently fewer volunteers, at the highest level is representative of the fact that this course is an elective even for college-bound students.[7]

In addition to the similarity of background of each participant, the instructional input that each student had received was quite parallel. Each student had used the same texts at each level and had the same instructors. Both instructors were nonnative speakers of Spanish who speak fluently and accurately but would not be mistaken for native speakers of Spanish.[8] In terms of language instruction, the copula contrast is presented quite early in the process of acquisition and is dealt with at each subsequent level in greater depth. Students had received explicit grammatical instruction on the copula contrast along with ample negative feedback on written work. Nevertheless, all instruction was based on traditional binary oppositions such as [±inherent characteristic] rather than on the interaction of such features. The sociolinguistic approach to copula contrast does not surface in the textbooks or in classroom discussion. Given the stages of SLA found in previous research, it is generally believed that *ser* is more frequent in the input learners receive, perhaps because of the descriptive tasks that generally occur in early language instruction, but no studies have quantified the frequency of each copula in a classroom setting (VanPatten, 1987).

Instruments

The Background Questionnaire. Although each student participated in the same learning context, a background questionnaire was administered to ensure that all participants had similar language backgrounds. The questionnaire investigated academic language-learning experience, foreign travel, contact with native Spanish (both in person and through the Internet), and language-learning success as indicated by course grades (both current and past). The data contributed by five participants were excluded from the analysis due to unique individual characteristics such as fluency in another language or participation in other language education programs, leaving one fewer participant at each level of enrollment except for level 3, from which the data from two participants were eliminated.

The Guided Interview. Each student participated in a tape-recorded guided conversation modeled after the sociolinguistic interviews described by Silva-Corvalán (1994) and Gutiérrez (1992). These interviews use a list of questions to guide each conversation but allow an interviewer to follow the conversation according to the interests of the participant. Questions included in the current study were aimed at eliciting descriptions in a natural way and were designed to avoid using copulas in the questions themselves. For example, students were instructed to describe a friend or relative with the command *Descríbele a tu hermano* "Describe (him) your brother" rather than

¿Cómo es? "What is he like?" (with *ser*). In this way, no information as to appropriate copula choice was provided. Questions were also designed to elicit comparisons between one person at two different points in time, thereby providing relevant data for the Frame of Reference variable. Each interview generally lasted between 10 and 20 minutes. Sample questions from the interview are included in Appendix A.

The Picture-Description Task. Each participant completed a picture-description activity in which they were asked to describe the people and events in each picture. In some cases, students were asked to describe a series of pictures that showed changes across time. For example, one series of pictures shows a couple growing older with each successive picture. The intent was to elicit a context in which an individual frame of reference was appropriate. The picture-description task is common in SLA research on copula choice because it serves to push participants to use certain adjectives (e.g., *muerto* "dead") that do not occur frequently in normal conversation (Briscoe, 1995; VanPatten, 1987). A total of nine pictures, grouped into three sets, were used. All participants described each of the pictures and were asked to provide more details when possible.

The Contextualized Questionnaire. Each participant completed a contextualized questionnaire.[9] Each item on the questionnaire contained a short paragraph that provided a discourse context that placed certain restrictions on the appropriate copula options. Each context was followed by two sentences that were identical except for the copula. Each participant selected the most appropriate response from three choices: (a) a sentence with *ser*, (b) the identical sentence with *estar*, or (c) that both sentences are appropriate. The order of presentation of each copula was varied randomly throughout the questionnaire, and the choice was worded as a preference for sentence A or B so that each item had the same response format. The contextualized questionnaire further provided situations in which one or more of the features said to influence copula choice are in conflict, thereby forcing learners to decide which constraint is more important in choosing a copula in the specified context. Examples of items from the contextualized questionnaire are shown in Appendix A.

Procedure

Each participant agreed to meet with the researcher during a free class period or before or after school in a study room in the school's library. Each participant completed the background questionnaire and secured parental permission prior to the interview time. Each session was tape-recorded and began with an explanation of how the microphone functioned. The researcher began with the interview task, followed by the picture-description task, and finally the contextualized questionnaire. This order was preferred for two reasons. First, the interview task serves to accustom students to the researcher and to

speaking in Spanish, as well as to put them at ease. Second, the questionnaire makes explicit the focus of the study and if this task were to precede the others, making the purpose of the descriptions explicit as well, participants would be inclined to focus more on their grammatical accuracy than on the meaning of their utterances. The entire session was completed within 50 minutes.

Procedures for Coding

The tape-recording of the interview and picture-description task were used to transcribe each session in its entirety prior to coding. Each instance of the syntactic structure [copula + adjective] that appeared in any of the three tasks was analyzed as an individual token in the current study, and all coding decisions were made with reference to the surrounding discourse context. All coding was done by the researcher. Additionally, one-tenth of the tokens were coded independently by a Spanish-speaking linguist. In cases of disparity, a consensus was reached through discussion of that particular token. In no case was it impossible to resolve the disagreement.

Each token was coded for three variables specific to the individual investigation: Response Type, Task Type, and Grammatical Accuracy. The variable Response Type included the categories *ser*, *estar*, omission, correct other (such as *parecer* "to seem"), and incorrect other. Task Type indicates whether the token appeared during the interview, the picture-description task, or in the questionnaire. The variable Grammatical Accuracy does not reflect accuracy of the copula choice but rather accuracy of features such as person, number, and tense. The categories within this variable include correct, omission, incorrect tense, incorrect number, and completely incorrect. This variable is not based on copula choice. Each copula + adjective token was also coded for the variables used in the description of the language-change data. Individual variables are listed in Table 3 along with the categories they entail; examples of each variable are provided from the data (no grammatical errors have been corrected).

There is one variable mentioned in studies of language change that was not included in the current analysis. In the process of applying the variable Attribute of Circumstance it became evident that this variable was problematic. The definition provided for it in sociolinguistic research is that it is an adjective that the speaker prefers to view as nondefining and thus circumstantial (Silva-Corvalán, 1986) or subjective (Gutiérrez, 1992). The most frequent example of this type of adjective is *calvo* "bald", in contexts where it is paired with *estar* to avoid the implication that it is an inherent characteristic. Although both Silva-Corvalán (1986, 1994) and Gutiérrez (1992) employed this feature to describe changing copula choice, they do not code their data specifically for this variable. Instead, the concept is employed in the analysis of specific examples. In the current study, it proved difficult to employ this notion as a variable for which all data could be coded because of the subjectivity

Table 3. Variables, categories, and examples of coding scheme

Variable	Examples from the data
Task type	
Guided interview	
Picture description	
Questionnaire	
Response type	
Ser	*Mi casa es pequeño* (B33) "My house is small"
Estar	*Cuando los niños estaba joven* (B103) "When the boys were young"
Correct other	*Parece bastante simpática* (I118) "(She) seems fairly nice"
Incorrect other	*Mira um . . . mira fuerte* (AA88) "(He) looks um he looks strong"
Omission	*Ellos simpáticos* (P88) "They nice"
Both	[Contextualized questionnaire only]
Grammatical accuracy	
Correct	*El muchacho es muy joven* (Q58) "The boy is very young"
Completely incorrect	*La familia eran muy divertido* (E91) "The family were very fun"
Incorrect tense	*Y mañana el hombre está enfermo* (I76) "And tomorrow the man is sick"
Incorrect number	*Zapatos es brown* (X84b) "Shoes is brown"
Omission	*Una bebe y muy alegre* (I83) "A baby and very happy"
Frame of reference	
Individual frame	*Durham hoy está muy diferente* (AE84) "Durham today is very different"
Class frame	*Es un van, es azul* (AK58) "It is a van, it is blue"
Susceptible to change	
Susceptible	*Tiene flores um están alegre* (AW119) "He has flowers um they are happy"
Not susceptible	*El padre es simpático* (AW74) "The father is nice"
Animacy	
Animate	*Está probablemente rica* (AS252) "She is probably rich"
Inanimate	*Es posible que no sabe* (AS259) "It is possible that she doesn't know"
Dependent on experience	
Dependent	*Ella no es muy alta* (AU70) "She (speaker's sister) is not very tall"
Not dependent	*La mujer es más gorda* (AU245) "The woman (in a photo) is fatter"
Semantic transparency	
Estar required	*Está muy sorprendido* (AA150) "She is very surprised"
Ser required	*Soy simpático* (BA14) "I am nice"
Modality contrast	*Está gordo* (BC133) "He is fat"
Meaning change	*Está enfermo* (AX100) "He is sick"
Near synonymy	*Están casados* (BM188) "They are married"
Adjective class	
Age	*Es un poquito más joven* (Q59) "He is a little younger"
Size	*Mi casa es bastante grande* (I71) "My house is fairly big"
Physical appearance	*Ella es, no es alto* (E22) "She is, she is not tall"
Description/evaluation	*La cultura es diferente* (L119) "The culture is different"

Table 3. (*Continued*)

Variable	Examples from the data
Description of a person-(ality)	*Son deportista y simpáticos* (B19) "They are sporty and nice"
Color	*Es blanco y azul* (G19) "It is white and blue"
Mental state	*El está aburrido* (P62) "He is bored"
Physical state	*El hombre está enfermo* (G76) "The man is sick"
Sensory characteristic	*La sopa es/está muy fría* (Questionnaire) "The soup is (*ser* or *estar*) very cold"
Status/class	*(Las personas) son blanco* (M79) "(The people) are white"
Miscellaneous	*Están juntos* (BI196b) "They are together"

inherent in the definition. For example, to state that a person is overweight may be viewed as circumstantial or as a defining characteristic. Short of a clear understanding of the speaker's intention—an impossibility with the current data—it was not possible to apply this feature.[10] Thus, the variable Attribute of Circumstance was eliminated from the analysis, and the comparison between the stages of SLA and language change will be based on Frame of Reference and Susceptibility to Change.

There are two features mentioned in sociolinguistic research that do serve to capture two more concrete aspects of circumstantiality. First, as is mentioned by Gutiérrez (1992, 1994b), these unflattering adjectives belong to a few specific adjective classes such as "description of a person(ality)" or "age" and, thus, the concept of a (potentially subjective) description of a referent is included in the analysis via an examination of the effects of the variable Adjective Class. Another important variable is Dependence on Experience, which is mentioned by Silva-Corvalán (1986, p. 590) as one of the possible dichotomies included in the bundle of features that Falk (1979) claimed interact to explain copula choice. Although Silva-Corvalán chose to use other variables in the description of the stages of change, the concept provides a nonsubjective means of categorizing data. Whether a speaker has prior experience with a referent that they are describing is not difficult to ascertain. In fact, in no case in the current data did this variable present difficulties. It should be pointed out that this variable is related to circumstantiality, given that all of the examples that Silva-Corvalán and Gutiérrez (1992) provided are situations in which the speaker has prior experience with the referent. Furthermore, despite an overlap, this variable is not identical to Frame of Reference because one may have prior experience with a referent and continue to compare that referent to a class of items (e.g., with *calvo* "bald"). Because of these facts, the variable Dependence on Experience was included in the current investigation but is not analyzed during the discussion of the stages of acquisition.

Statistical Analysis

The dependent variable for all statistical tests in the current study was the learner response. Although this variable was coded for all possible responses (e.g., omission), these categories were collapsed into two for the purposes of analysis, making the two categories for the variable Response Type "*estar*" and "not *estar*". Thus, each statistical test examines the effect of a particular variable or variables on the use of *estar*. The motivation for examining the effect of each variable on the use of *estar* rather than *ser* is twofold. First, in sociolinguistic research, it is *estar* that is shown to have changing constraints, rather than *ser* (Silva-Corvalán, 1986). Second, in the research on the SLA of copula choice it was shown that the initial stages of development exhibit overgeneralization and exclusive use of *ser* (VanPatten, 1987). Because *estar* is the copula that is slower to emerge, it is the contextual features that predict its emergence that are of interest. This approach differs from studies of the acquisition of copula choice in Spanish based on an error-analysis design (Briscoe, 1995; Ramírez-Gelpi, 1995; Ryan & Lafford, 1992; VanPatten, 1985, 1987) because the current investigation avoids the potentially problematic practice of evaluating learner accuracy even when a speaker's intentions are unknown.[11] Instead, the relationship between SLA and language change is examined through an analysis of the contextual variables that could potentially influence copula choice.

To examine the effects of those variables found to describe the stages of language change (viz. Frame of Reference and Susceptibility to Change), each variable was submitted to an individual chi-square test. These tests examine the strength of the correlation of each contextual variable with the usage of *estar*. Because Silva-Corvalán (1986, 1994) discussed the effects of particular variables, such as Semantic Transparency, in terms of chi-square results computed in the SPSS crosstabs program, this test is a reasonable starting point for the discussion of the current hypothesis. Silva-Corvalán coded data for innovative uses of *estar* but not for other contextual features, and the stages she set forth are not based on statistical analyses. Nevertheless, a significant effect of the variable Susceptibility to Change on the use of *estar* at early and all subsequent levels of enrollment, and a significant effect of the variable Frame of Reference only at higher levels of enrollment, would provide comparable results to those presented for language change.

In addition to the use of individual statistical tests to determine the strength of relationship between *estar* and a particular feature, a logistic regression analysis was employed. Logistic regression is an appropriate statistical test for linguistic data in general because it does not require that the data be normally distributed (see Berdan, 1996, for a description of logistic regressions and their application to SLA data).[12] This test allows for categorical variables of the type described for the current investigation and is able to deal with variables that are not binary. A stepwise regression analysis of the type employed in the current investigation gradually selects those factors that best

Table 4. Total tokens and percentage of total tokens for response type

Response	Level							
	1		2		3		4	
Ser	654	(62.9)	859	(68.8)	979	(65)	323	(63.5)
Estar	215	(20.7)	244	(19.5)	370	(24.6)	132	(25.9)
Correct other	5	(0.4)	6	(4.8)	55	(3.7)	26	(5.1)
Incorrect other	17	(1.6)	25	(2)	29	(1.9)	7	(1.4)
Omission	138	(13.3)	104	(8.3)	58	(3.9)	18	(3.5)
Both	10	(1.0)	11	(0.8)	13	(0.9)	3	(0.6)
Total	1039	(100)	1249	(100)	1504	(100)	509	(100)

Note. Numbers in parentheses represent percentages.

predict the appearance of the dependent variable—in this case, *estar*. With each step, a new variable is added to the equation, thereby increasing the predictive strength of the model. The analysis continues until all variables have been added to the equation that can make a significant contribution to the prediction of *estar*. Thus, the discussion of the results of these tests rests on those variables that are significant predictors of the use of *estar* when considered in conjunction with all other variables.

The inclusion of this additional statistical test is necessary because, as Falk (1979) and Silva-Corvalán (1994) have shown, the use of either copula is based on the interaction of a number of contextual features. Because features that predict the use of *estar* are generally aligned, each time *estar* appears it is likely that most of the contextual features are also present. This implies that an individual correlation test, such as a chi-square, will show that a particular feature is generally present when *estar* is used. It tells nothing about whether that feature is actually a good predictor of learner usage of *estar*. Only in cases in which the contextual features conflict can one feature be shown to be more salient than another or more relevant to a learner grammar at a particular point in time. Furthermore, the inclusion of all of the features associated with sociolinguistic descriptions of copula choice allows for a more complete discussion of the comparison between the process of language acquisition and that of language loss while at the same time providing a more closely guided investigation than that provided in Geeslin (2000).

RESULTS

A total of 4301 tokens were collected, the distribution of which is reflected in Table 4, with the total number of tokens of each type followed by the percentage of the total that this number represents. An initial assessment of the results reveals that *ser* is used much more frequently than *estar* at each level. The figures for the percentages of each type of response further imply that use of *estar* increases slightly across levels and the omission of copulas de-

Table 5.   Total tokens of *ser* and *estar* for susceptibility to change ([±change])

	+Change			−Change			Total tokens
Level	*Ser*	*Estar*	Total	*Ser*	*Estar*	Total	
1	388 (70.93)	159 (29.06)	547	436 (88.62)	56 (11.38)	492	1,039
2	531 (75.11)	176 (24.89)	707	474 (87.45)	68 (13.82)	542	1,249
3	507 (64.67)	277 (35.33)	784	627 (87.08)	93 (12.92)	720	1,504
4	176 (63.77)	100 (36.23)	276	201 (86.27)	32 (13.73)	233	509

Note. Numbers in parentheses represent percentages.

Table 6.   Total tokens of *ser* and *estar* for frame of reference ([±individual])

	+Individual			−Individual			Total tokens
Level	*Ser*	*Estar*	Total	*Ser*	*Estar*	Total	
1	211 (57.81)	154 (42.19)	365	613 (90.95)	61 (9.05)	674	1,039
2	206 (52.82)	184 (47.18)	390	799 (93.02)	60 (6.98)	859	1,249
3	230 (47.62)	253 (52.38)	483	904 (88.54)	117 (11.46)	1,021	1,504
4	71 (43.83)	91 (56.17)	162	306 (88.18)	41 (11.82)	347	509

Note. Numbers in parentheses represent percentages.

creases. This is consistent with the results of previous studies of the SLA of copula choice.

Table 5 provides numerical information about the use of *estar* in contexts that are [+susceptible to change] in contrast with those that are not. Recall that *estar* is generally paired with contexts where, given a particular referent, the attribute with which it is paired is considered to be susceptible to change. It would not be accurate, however, to claim that standard use would be 100% because there are a number of features that interact to determine copula choice. Totals and the percentage of the total tokens this represents are provided in Table 5 for the usage of both *ser* and *estar* as a point of comparison. The number of tokens belonging to each category suggests that the number of contexts where the attribute is susceptible to change is relatively equal to the number of contexts where it is not, and this distribution does not vary drastically from one level to the next. Additionally, it appears that *estar* is used more often with contexts that are [+susceptible to change], just as it is in language-change data. Finally, the percentage of use of *estar* in contexts that are [+susceptible to change] appears to increase slightly as length of study increases, but because *estar* is actually used more frequently at the first level than at the second it will be necessary to investigate this trend more closely.

Table 6 shows the numerical breakdown of copula choice for all tokens grouped according to the variable Frame of Reference. The categorization

Table 7. Results of a chi-square test on susceptibility to change and *estar* at all levels of enrollment

Level	df	Number of tokens	Value of X^2
1	1	1039	48.30*
2	1	1249	28.98*
3	1	1504	100.46*
4	1	509	32.13*

*$p < .05$.

Table 8. Results of a chi-square test on frame of reference and *estar* at all levels of enrollment

Level	df	Number of tokens	Value of X^2
1	1	1039	158.47*
2	1	1249	273.13*
3	1	1504	293.81*
4	1	509	110.83*

*$p < .05$.

[±individual] distinguishes the individual frame of reference (comparison to the referent) from the class frame of reference (comparison to a group). Recall that contexts that are [+individual] are generally paired with *estar*. Again, standard use would not be 100% appearance of *estar* because other variables interact to determine copula choice. The number of tokens of both *ser* and *estar* that appeared in each of the two contexts for each of the four levels of enrollment is provided in Table 6, along with the percentages that each total represents. As with contexts that are susceptible to change, the tabulations show that *estar* is used more frequently with the individual frame of reference. This is the expected result for both SLA data and for language-change data. Additionally, the frequency with which *estar* appears increases with the level of enrollment. One interesting difference between the two variables is that [+individual] appears only about one third of the time, with the [−individual] (class) frame of reference accounting for all other contexts. Nevertheless, this too is relatively consistent across levels.

Tables 5 and 6 show, as expected, that *estar* corresponds to the [+change] and the [+individual] categories just as with native speech. Nonetheless, a chi-square test was performed for each variable at each level of enrollment to ensure that this relationship was significant. Tables 7 and 8 show the results of the individual chi-square tests for the effects of Susceptibility to Change and Frame of Reference on the appearance of *estar*. There is only one degree of freedom for each contextual variable because only two categories are available. The number of tokens and the value of the chi-square are provided for

Table 9. Constraints on usage of *estar* based on a regression analysis of sociolinguistic variables

| | Level | | | |
Constraint	1	2	3	4
Grammatical accuracy	*X	X	X	X
Adjective class	*X	*X	*X	*X
Semantic transparency		*X	*X	*X
Animacy				
Susceptibility to change	*X	*X	*X	*X
Dependence on experience	*X	*X	*X	*X
Frame of reference		X	*X	*X
Task	*X	*X	*X	*X

*$p < .05$.

each level of enrollment. A *p*-value smaller than .05 was considered significant. The results show a statistically significant relationship at all levels of enrollment between the variables Susceptibility to Change and Frame of Reference and the use of *estar*.

Given the results of the two chi-square tests shown in Tables 7 and 8, it can be claimed that each of the two variables used to describe the stages of language loss have been shown to be related to the appearance of *estar* in learner data as well. Nevertheless, it remains unclear what stages of acquisition based on these features can be proposed. This is because the chi-square tests show significant results at each level of enrollment and because they do not take into account the interaction of many contextual features at one time. It may well be the case that a single feature that is correlated with the appearance of *estar* is not the best predictor of its use at that level of development nor the best descriptor of changes in learner grammar across time. To evaluate the overall importance of each factor in predicting the appearance of *estar*, a forward stepwise regression test including each of the contextual variables employed in sociolinguistic studies was performed for each level of enrollment. Unlike Geeslin (2000), only those contextual variables related to sociolinguistic descriptions of copula choice are included in the analysis. The results from each of the four tests, one for each level of enrollment, provide an equation that includes only those features that are significant predictors of the usage of *estar* at that level. As a result, it is possible to compare those features that are most relevant for the selection of *estar* at a particular level of development to those that are described in Silva-Corvalán's (1986) stages of language change.

Table 9 is a compilation of the results from each of the four regression tests. A total of eight variables were included in these tests. For each contextual variable at each level of enrollment, an X signifies that this factor was included in the equation used to predict the appearance of *estar*. An alpha

level of .05 was used, and those factors that were significant predictors of *estar* within the equation are indicated with an asterisk. Because a stepwise regression adds one variable at a time to the equation, it is possible that a variable has a significant *p*-value at the time it is added to the equation but that, upon adding other variables at later stages, the significance of those added earlier decreases. In some cases, the variable is eliminated from the equation at later stages, but in others it continues to contribute to the overall equation despite the fact that the *p*-value rises above .05. The complete results from each individual regression are included in Appendix B.

Like the results of the chi-square tests, the regression analysis shows that the variable Susceptibility to Change is related to the usage of *estar* at all levels, even when considered with other contextual features. In contrast, the variable Frame of Reference, also shown to be statistically related to the usage of *estar* at all levels when considered in isolation, is a significant predictor of its use only at the higher levels of enrollment. This variable was completely excluded from the equation at level 1. At the next level of enrollment, Frame of Reference was the first variable added to the regression equation, but with each progressive cycle it became less significant, arriving at a *p*-value of .075 in the final equation. At levels 3 and 4, this variable was included in the equation and it reached significance. Additionally, the results of the regression analyses show that other sociolinguistic variables such as Adjective Class and Dependence on Experience are important descriptors of learner usage of *estar*.

DISCUSSION: SLA AND LANGUAGE CHANGE

As stated previously, our purpose was to test the hypothesis that the stages of SLA exhibited in the present study, based on the variables Frame of Reference and Susceptibility to Change, would mirror the stages set forth for language change by Silva-Corvalán (1986). Silva-Corvalán showed that in contexts of language loss, certain constraints are lost before others as the grammar becomes less restricted in terms of permissible contexts for *estar*. First, the frame of reference constraint was lost, followed by a loss of the restriction susceptibility to change, followed by contexts where no restrictions were placed on the usage of *estar* with adjectives. Thus, given the hypothesis, the prediction for SLA data is that learners will acquire the restriction susceptibility to change earlier in the acquisition of copula contrast than the frame of reference restriction. The results of the regression analysis appear to support this claim. At all levels of enrollment included in the current study, the variable Susceptibility to Change is a significant predictor of the appearance of *estar*. In contrast, the variable Frame of Reference exhibits development over time from complete exclusion from the regression equation at level 1 to inclusion and a significant *p*-level at the two highest levels of enrollment. Thus, those constraints lost during the process of the extension of *estar* are those acquired during the process of SLA. This relationship is summarized in Table 10.

Table 10. Comparison of constraints in SLA and language change

Language process	Susceptibility to change	Frame of reference
SLA	Learned early	Learned late
Change/loss	Lost late	Lost early

The investigation of the mirror-image hypothesis allows an examination of those processes found in the current investigation to represent SLA. For example, it is interesting to consider why the susceptibility to change constraint is acquired earlier than the frame of reference constraint. Most likely, there is a relationship between the frequency of the context in the input, the degree to which the constraint is taught, and the significance of that constraint in the prediction of the appearance of *estar*. Because contexts that are [+susceptible to change] account for about half of the contexts in the current study, it is likely that in-class conversations between students also contain a higher number of such contexts than those where the [+individual] frame of reference is relevant. Furthermore, many of the lexical items frequently paired with *estar* that are taught and learned very early, such as *alegre* "happy" and *enojado* "angry," can be explained by susceptibility to change. Also, most textbooks, and indeed the one used by the participants in the current investigation, directly address this constraint. It can be said, then, that learners may begin with constraints such as susceptibility to change and modify these only on exposure to a significant number of counterexamples. Such counterexamples, as in the use of *estar* with *guapo* "handsome" to express an individual frame of reference and indicate that a person "looks great," can be explained by the Frame of Reference variable even though the Susceptibility to Change variable would predict that one's physical appearance is a permanent characteristic. It is probable that these stages of development represent the process through which language learners specify grammatical rules once the initial generalization is no longer sufficient. In this sense, the investigation of the current hypothesis has provided insights into the process of SLA that would not otherwise have been examined.

Nevertheless, although the simplicity of the mirror-image hypothesis is attractive, results from the current investigation suggest that the relationship between language change and SLA is more complex. For example, the two variables used in the current investigation are those that Silva-Corvalán (1986) believed would best categorize her data. Nevertheless, she did not show (nor aim to show) that other contextual variables are any less effective. Because the results of the current investigation show that there are other contextual features that are relevant in predicting the appearance of *estar* in learner data, it is reasonable to wonder how these features compare across disciplines. For example, dependence on experience and adjective class show results identical to those for susceptibility to change in that each is a significant predictor of

the use of *estar* for all levels of enrollment. Given this fact, it is possible to claim that any one of these three variables is acquired before the frame of reference constraint. Additionally, semantic transparency displays a similar pattern of acquisition to that of frame of reference because it too is significant only at higher levels of enrollment. In combination, these five variables provide a number of potential hypotheses that could be tested on data from language change. Furthermore, the fact that these additional variables are not included in the mirror-image hypothesis does not diminish the importance of the additional information they provide about learner development. To limit the discussion of SLA to the mirror-image hypothesis would be to overlook these insights.

Another important consideration for the mirror-image hypothesis is that it entails only the order in which constraints are lost or acquired. Whereas it is assumed that the endpoint of SLA is similar to the beginning point of language loss, the same does not hold true for the endpoint of language loss and the beginning point of SLA. Instead, the beginning stages of SLA show copula omission or overgeneralization of *ser* and a complete lack of *estar*, whereas the endpoint for language loss exhibits a complete extension of *estar* to all contexts. The difference between these two processes can be seen as a difference in the simplified grammar from which, or toward which, a speaker progresses. These differences do nothing to detract from the usefulness of the mirror-image hypothesis as a tool for the examination of the SLA of copula choice, but they do constrain the extent to which this approach can be applied.

Despite the limitations of the mirror-image hypothesis, it has a very important advantage. The current investigation shows a principled examination of SLA data based on those features found to be valuable descriptors of copula choice in sociolinguistic research. The application of these features speaks to the importance of cross-disciplinary investigations and the application of tools from one field of inquiry to data from another. Furthermore, the application of only those features from sociolinguistic research to the SLA data shows a major improvement over Geeslin (2000) for two reasons. First, the current study is guided by the examination of the mirror-image hypothesis in such a way that the analysis is based on the investigation of a single point of comparison of SLA data to sociolinguistic data. Second, those features included in the analysis are clearly related to one another and have been previously applied to data in unison. In contrast, Geeslin (2000) showed that copula choice must be examined in terms of multiple interacting contextual features, but those features lacked a unifying theoretical approach and, thus, many of the features raised questions that were unanswerable. This improvement is demonstrated by the fact that no variable in the current investigation showed a failure to contribute to the description of acquisition across time. Instead, all variables are either present from the earliest level of enrollment and continue to be significant predictors at all subsequent levels or the variables show evidence of gaining significance as time progresses. In no case

does a variable show significance at only one level of acquisition nor does it show intermittent significance across levels.

CONCLUSIONS AND FUTURE DIRECTIONS

The current analysis of SLA data has shown that the description of the stages of language loss (Silva-Corvalán, 1986, 1994), based on the variables Susceptibility to Change and Frame of Reference, provides a valuable tool for the analysis of the SLA of copula choice. Through an investigation of the mirror-image hypothesis predicted to hold between these two processes, it was found that the frame of reference constraint that is lost earlier in the process of language change is also acquired later in the process of SLA while, at the same time, the susceptible to change constraint is not only lost late but is acquired early. Additionally, many of the contextual features used to describe language change that were not included in the stages themselves, such as adjective class and semantic transparency, also provided additional information about the process of acquiring copula choice in Spanish. These results underscore the potential of cross-disciplinary research to strengthen investigations in SLA and at the same time highlight new directions for sociolinguistic research. For example, the results from the current investigation suggest that it would be useful to apply the coding scheme from the current investigation to sociolinguistic data to investigate the interaction of each of the contextual features and to further specify the hypotheses for the relationship that exists between language acquisition and language loss.

The current investigation also signals the need for future investigations of copula choice. To date, there exists no in-depth investigation of the influence of a single contextual feature on language development. Despite the fact that adjective class is a significant predictor of learner behavior at all levels of enrollment, the current analysis does not address each individual adjective class in terms of its effect on learner use of *estar*. The feature semantic transparency shows an increased importance as acquisition proceeds, and it would be interesting to examine each individual category of this variable in terms of its contribution to predicting learner use of *estar*. Finally, there exists no research on more advanced learners nor on native speakers with which the current study can be compared. These individual research questions represent directions for future inquiry, each of which relies on a principled, interdisciplinary approach to second language data.

(*Received 13 August 2001*)

Notes

1. Although some periphrastic forms may remain in languages today rather than having moved on to the development of grammaticalized bound morphemes, both language change and language acquisition are seen here to move through similar stages. Giacalone Ramat (1995) explained that both are driven by a preference for analytical forms that are more transparent than synthetic ones.

The fact that French has maintained the periphrastic perfect instead of the preterit shows that there are other factors that can interact with internal language processes and can influence the final outcome of a change.

2. Although Silva-Corvalán (1986, 1990) was able to show parallels between language systems undergoing simplification, she cautioned that one should always take into account the social setting, where both internal and external forces are at work.

3. This hypothesis, though appropriate for the examination of copula choice, is not applicable to all language change, given that some changes involve an increase in complexity (see Preston, 1989, for a discussion of change from above).

4. De Jonge (1993) also found that *estar* was more frequently used when the word *ya* "already" was present and that the frequency of innovation was greater in Caracas.

5. Ryan and Lafford (1992) also found a stage at which *estar* was overgeneralized, but this dip in accuracy of *ser* is likely to be consistent with other SLA research. This is because the use of an error analysis that relies on a minimum percentage of correct tokens to determine each stage may obscure changes in learner speech patterns if those changes do not cause accuracy rates to fall below the minimum score.

6. The data set analyzed in Geeslin (2000) is the same as that analyzed in the current investigation and will be described in greater depth along with the methods for the current study. The research questions, analysis of the data, and results are distinct.

7. It was the goal of the researcher to have a minimum of 20 participants at each level of enrollment and, for the first three levels of enrollment, recruiting ended as soon as 20 participants had volunteered, although additional volunteers were accepted. At the highest level, level 4, only one student chose not to participate in the current investigation. Nevertheless, low student enrollment made it impossible to include more students at this level in the current investigation.

8. Although one of the instructors had studied in Mexico, the time of study was more than 25 years ago. Both instructors travel to Spain frequently and often lead student trips to Spain.

9. The use of the term "questionnaire" is employed here both to maintain consistency with Gutiérrez (1992) and Silva-Corvalán (1994), who refer to a fill-in questionnaire, and to indicate the variable nature of the content under investigation. Although a grammaticality judgment task implies a prescriptive ("correct") answer, a questionnaire is generally employed to survey individual preferences. Given the changing status of copula choice, the term "questionnaire" was deemed appropriate.

10. Although this variable is criticized in terms of the current investigation, I believe that Attribute of Circumstance would be problematic for sociolinguistic data as well. The view that an adjective is "unfortunate" or "unflattering" and that a speaker chooses not to define the referent in terms of this adjective may in fact be the perception of the investigator rather than the speaker. It should be noted that neither Silva-Corvalán (1986) nor Gutiérrez (1992) actually coded data for this variable. Instead, the notion of circumstance is used to facilitate discussion of their results.

11. In a more recent investigation, Geeslin (2001) addressed learner accuracy in contexts in which native speaker choice varies without relying on the error-analysis design employed in earlier studies. This was accomplished by a comparison of the questionnaire data from the 72 participants included in the current study and native speaker responses on the same contextualized questionnaire ($n = 10$). Geeslin (2001) differs significantly from the current investigation because only those questionnaire contexts where native speakers unanimously selected the same copula were included in the analysis. Moreover, whereas the current investigation tests the mirror-image hypothesis, Geeslin (2001) was specifically designed to address learner accuracy. It was found that those items for which all native speakers selected *ser* showed no evidence of improvement across time because of the tendency to overgeneralize *ser* at early stages of acquisition. In contrast, those contexts for which all native speakers selected *estar* showed a significant relationship between length of time studying Spanish and accuracy.

12. The statistical package with which all statistical tests were conducted was SPSS version 10.1 for PC. This package was initially selected in an effort to maintain consistency with other studies on copula choice (Silva-Corvalán, 1986, 1994) and so that more than one statistical test could be performed within a single package.

References

Adamson, H., & Regan, V. (1991). The acquisition of community norms by Asian immigrants learning English as a second language: A preliminary study. *Studies in Second Language Acquisition*, *13*, 1–22.

Andersen, R. (1979). Expanding Schumann's pidginization hypothesis. *Language Learning*, *29*, 105–119.

Andersen, R. (1981). Two perspectives on pidginization as second language acquisition. In R. Andersen (Ed.), *New dimensions in second language acquisition research* (pp. 165–196). Rowley, MA: Newbury House.

Berdan, R. (1996). Disentangling language acquisition from language variation. In R. Bayley & D. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 203–244). Amsterdam: Benjamins.

Bickerton, D., & Odo, C. (1976). *Change and variation in Hawaiian English: Vol. 1. General phonology and pidgin syntax*. Honolulu: Social Sciences and Linguistics Institute, University of Hawai'i.

Briscoe, G. (1995). *The acquisition of* ser *and* estar *by non-native speakers of Spanish*. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Clements, J. C. (1988). The semantics and pragmatics of the Spanish <copula + adjective> construction. *Linguistics*, *26*, 779–822.

De Jonge, B. (1993). (Dis)continuity in language change: *Ser* and *estar* + age in Latin American Spanish. In F. Drijkoningen & K. Hengeveld (Eds.), *Linguistics in the Netherlands* (pp. 69–80). Amsterdam: Benjamins.

Dorian, N. (1980). Maintenance and loss of same-meaning structures in language death. *Word*, *31*, 39–45.

Falk, J. (1979). *Ser y estar con atributos adjetivales* [*Ser* and *estar* with attributive adjectives]. Stockholm: Alqvist & Wiksell.

Geeslin, K. L. (2000). A new approach to the second language acquisition of copula choice in Spanish. In R. Leow & C. Sanz (Eds.), *Spanish applied linguistics at the turn of the millennium: Papers from the 1999 Conference on the L1 and L2 Acquisition of Spanish and Portuguese* (pp. 50–66). Somerville, MA: Cascadilla Press.

Geeslin, K. L. (2001). Changing norms, moving targets, and the SLA of copula choice. *Spanish Applied Linguistics*, *5*, 29–55.

Giacalone Ramat, A. (1995). Iconicity in grammaticalization processes. In R. Simone (Ed.), *Iconicity in language* (pp. 119–139). Amsterdam: Benjamins.

Greenberg, J. (1991). Typology/universals and second language acquisition. In T. Huebner & C. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theories* (pp. 3–22). Amsterdam: Benjamins.

Gunterman, G. (1992). An analysis of interlanguage development over time: Part 2. *Ser* and *estar*. *Hispania*, *75*, 1294–1303.

Gutiérrez, M. (1992). The extension of *estar*: A linguistic change in progress in the Spanish of Morelia, Mexico. *Hispanic Linguistics*, *5*, 109–141.

Gutiérrez, M. (1994a). La influencia de "los de abajo" en tres procesos de cambio lingüístico en el español de Morelia, Michoacán [The influence of "those from below" in three processes of linguistic change in the Spanish of Morelia, Michoacán]. *Language Problems and Language Planning*, *18*, 257–269.

Gutiérrez, M. (1994b). Simplification, transfer, and convergence. *Bilingual Review*, *19*, 111–121.

Hawkins, J. (1991). Language universals in relation to acquisition and change: A tribute to Roman Jakobson. In L. Waugh & S. Rudy (Eds.), *New vistas in grammar: Invariance and variation* (pp. 473–493). Amsterdam: Benjamins.

Lema, J. (1995). Distinguishing copular and aspectual auxiliaries: Spanish *ser* and *estar*. In J. Amastae, G. Goodall, M. Montalbetti, & M. Phinney (Eds.), *Contemporary research in Romance linguistics* (pp. 257–274). Amsterdam: Benjamins.

Luján, M. (1981). The Spanish copulas as aspectual indicators. *Lingua*, *54*, 165–210.

Milroy, J. (1992). *Linguistic variation and change: On the historical sociolinguistics of English*. Oxford: Blackwell.

Mufwene, S. (1990). Transfer and the substrate hypothesis in creolistics. *Studies in Second Language Acquisition*, *12*, 1–23.

Myhill, J. (1991). Typological text analysis: Tense and aspect in creoles and second languages. In T. Huebner & C. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theories* (pp. 93–122). Amsterdam: Benjamins.

Odlin, T. (1992). Transferability and linguistic substrates. *Second Language Research*, *8*, 171–202.

Preston, D. (1989). *Sociolinguistics and second language acquisition*. Oxford: Blackwell.

Preston, D. (1993). Variation linguistics and L2 acquisition. *Second Language Research*, *9*, 153–172.

Preston, D. (1996). *Variationist perspectives on second language acquisition*. In R. Bayley & D. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 1–45). Amsterdam: Benjamins.

Ramírez-Gelpi, A. (1995). *The acquisition of* ser *and* estar *among adult native English speakers learning Spanish as a second language*. Unpublished doctoral dissertation, University of Southern California, Los Angeles.

Romaine, S. (1988). Contributions from pidgin and creole studies to a sociolinguistic theory of language change. *International Journal of the Sociology of Language*, *71*, 59–66.

Romaine, S. (1994). *Language in society: An introduction to sociolinguistics*. Oxford: Oxford University Press.

Ryan, J., & Lafford, B. (1992). The acquisition of lexical meaning in a study abroad environment: *Ser + estar* and the Granada experience. *Hispania*, *75*, 714–722.

Sanz, C., & González, M. (1995). *Ser* and *estar* in Tortosí Catalan: Language contact, language variation, and language change. *Sintagma*, *7*, 5–25.

Schmitt, C. (1992). *Ser* and *estar*: A matter of aspect. *Proceedings of NELS*, *22*, 411–425.

Schumann, J. (1978). *The pidginization process: A model for second language acquisition*. Rowley, MA: Newbury House.

Silva-Corvalán, C. (1986). Bilingualism and language change: The extension of *estar* in Los Angeles Spanish. *Language*, *62*, 587–608.

Silva-Corvalán, C. (1990). Current issues in studies of language contact. *Hispania*, *73*, 162–176.

Silva-Corvalán, C. (1991). Cross-generational bilingualism: Theoretical implications of language attrition. In T. Huebner & C. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theories* (pp. 325–346). Amsterdam: Benjamins.

Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.

Smith, C. (1997). *The parameter of aspect* (2nd ed.). Dordrecht: Kluwer.

Solé, Y., & Solé, C. (1977). *Modern Spanish syntax: A study in contrast*. Boston: Heath.

VanPatten, B. (1985). The acquisition of *ser* and *estar* in adult second language learners: A preliminary investigation of transitional stages of competence. *Hispania*, *68*, 399–406.

VanPatten, B. (1987). The acquisition of *ser* and *estar*: Accounting for developmental patterns. In B. VanPatten, T. Dvorak, & J. Lee (Eds.), *Foreign language learning: A research perspective* (pp. 61–75). Rowley, MA: Newbury House.

Williams, J. (1989). Language acquisition, language contact, and nativized varieties of English. *RELC Journal*, *20*, 39–69.

APPENDIX A

TESTING MATERIALS

Sample Questions from the Guided Interview

1. *¿Cómo te llamas?* "What is your name?"
2. *¿Cuántos años tienes?* "How old are you?"
3. *¿Cuántas personas hay en tu familia?* "How many people are there in your family?"
4. *¿Cómo se llaman?* "What are their names?"
5. *Descríbele a tu mamá (tu papá, etc.).* "Describe your mother (your father, etc.)."
6. *Describe tu casa (tu coche).* "Describe your house (your car)."

Sample Items from the Contextualized Questionnaire

Instructions: Pretend you are a person who lives in Mexico. You will read descriptions below of situations that have taken place between your housemates Paula and Raúl.

Read the scenes below and decide which sentence of Raúl's you prefer. Please check only one option. Also, indicate how sure you are of the answer.

1. Paula and Raúl are planning to go out for a bite to eat. Paula is yelling from the bedroom while she gets ready in order to make plans with Raúl. As she comes out of her room she asks:

Paula: *¿Quieres ir en mi coche?* "Do you want to go in my car?"
A. Raúl: *¡Ay! ¡Qué guapa estás!*   "Wow! How pretty you are!"
B. Raúl: *¡Ay! ¡Qué guapa eres!*

_____ I prefer sentence A.
_____ I prefer sentence B.
_____ I like both A and B.

How sure are you?
     5   4   3   2   1
Very sure   Not sure at all

2. Paula and Raúl go to a local restaurant. They eat there frequently and the people who work there are always very nice. This time, Raúl has ordered something new on the menu and Paula is curious about what Raúl thinks of his meal:

Paula: *Raúl, ¿te gusta la comida?*   "Raul, do you like the meal?"
A. Raúl: *Sí, la cena es muy buena.*   "Yes, the dinner is very good."
B. Raúl: *Sí, la cena está muy buena.*

_____ I prefer sentence A.
_____ I prefer sentence B.
_____ I like both A and B.

How sure are you?
     5   4   3   2   1
Very sure   Not sure at all

3. Paula knows that Raúl is doing poorly in school. Raúl's father is working hard to pay for school and Raúl is afraid to tell him his grades. Paula wonders if he can keep his grades a secret:

Paula: *¿Tienes que hablar con tu papá?*   "Do you have to talk with your dad?"

A. Raúl: *Sí, claro, mi papá no está estupido.*   "Yes, of course, my dad is not stupid."

B. Raúl: *Sí, claro, mi papá no es estupido.*

_____ I prefer sentence A.
_____ I prefer sentence B.
_____ I like both A and B.

How sure are you?
     5   4   3   2   1
Very sure   Not sure at all

APPENDIX B

RESULTS FROM REGRESSION ANALYSES

Table B1. Significant predictors of *estar* at level 1

Variable	Coefficient (B)	*SE*	*df*	Significance	Estimated odds ratio
Task			2	.00	
Task (1)	.62	.62	1	.32	1.86
Task (2)	−1.91	.60	1	.00	.15
Accuracy			4	.00	
Acc(1)	8.91	54.21	1	.87	7435.83
Acc(2)	−12.04	164.26	1	.94	.00
Acc(3)	−2.77	.54	1	.00	.06
Acc(4)	9.25	12.13	1	.45	10355.50
Adjective			9	.00	
Adj(1)	−.02	.592	1	.98	.99
Adj(2)	.68	.455	1	.14	1.97
Adj(3)	.30	.526	1	.57	1.35
Adj(4)	.09	.45	1	.84	1.10
Adj(5)	−.27	.45	1	.54	.76
Adj(6)	−1.65	.44	1	.00	.19
Adj(7)	−3.17	.47	1	.00	.04
Adj(8)	1.23	.56	1	.03	3.41
Adj(9)	−1.11	.52	1	.03	.33
Dependent on Experience	1.62	.40	1	.00	5.05
Susceptible to Change	1.47	.36	1	.00	4.36
Constant	−1.87	1.09	1	.09	.15

Note. −2 Log likelihood = 650.203; model X^2 = 409.286; *df* = 17; *p* < 0.001.

Table B2.   Significant predictors of *estar* at level 2

Variable	Coefficient (B)	SE	df	Significance	Estimated odds ratio
Task			2	.00	
Task (1)	1.80	.54	1	.00	6.06
Task (2)	−1.33	.49	1	.01	.27
Accuracy			3	.78	
Acc(1)	8.49	70.07	1	.90	4852.18
Acc(2)	.17	.90	1	.85	1.19
Acc(3)	8.24	8.08	1	.31	3781.56
Semantic Transparency			4	.02	
Sem(1)	−.42	.72	1	.56	.66
Sem(2)	−.67	.54	1	.22	.51
Sem(3)	1.57	.56	1	.01	4.80
Sem(4)	−.43	.58	1	.46	.65
Adjective			9	.00	
Adj(1)	.26	.56	1	.64	1.30
Adj(2)	.51	.44	1	.24	1.67
Adj(3)	1.31	.55	1	02	3.71
Adj(4)	1.23	.78	1	.11	3.42
Adj(5)	.09	.76	1	.90	1.10
Adj(6)	−.75	.45	1	.10	.47
Adj(7)	−1.37	.66	1	.04	.25
Adj(8)	.89	.64	1	.17	2.42
Adj(9)	−.28	.54	1	.60	.76
Dependent on Exper-ience	2.67	.44	1	.00	14.33
Susceptible to Change	1.26	.45	1	.01	3.51
Frame of Reference	.47	.26	1	.08	1.59
Constant	−4.84	1.07	1	.00	.01

Note. −2 Log likelihood = 730.641; model X^2 = 503.114; *df* = 21; *p* < 0.001.

Table B3. Significant predictors of *estar* at level 3

Variable	Coefficient (B)	SE	df	Significance	Estimated odds ratio
Task			2	.00	
Task (1)	.73	.34	1	.03	2.07
Task (2)	−1.43	.33	1	.00	.24
Accuracy			5	.61	
Acc(1)	−11.17	60.43	1	.85	.00
Acc(2)	−2.30	1.6	1	.15	.10
Acc(3)	−.36	.82	1	.67	.70
Acc(4)	8.68	7.60	1	.25	5893.86
Acc(5)	5.70	60.43	1	.93	298.85
Semantic Transparency			4	.00	
Sem(1)	.25	.39	1	.53	1.28
Sem(2)	−1.10	.44	1	.01	.33
Sem(3)	−.51	.32	1	.11	.60
Sem(4)	−2.28	.47	1	.00	.10
Adjective			10	.00	
Adj(1)	.17	.46	1	.71	1.19
Adj(2)	.01	.42	1	.98	1.01
Adj(3)	.07	.41	1	.86	1.07
Adj(4)	.39	.51	1	.45	1.47
Adj(5)	−1.091	.49	1	.02	.34
Adj(6)	−1.17	.38	1	.00	.31
Adj(7)	−1.30	.46	1	.00	.27
Adj(8)	−.04	.56	1	.95	.97
Adj(9)	−.48	.47	1	.31	.62
Adj(10)	−.94	1.57	1	.55	.39
Dependent on Experience	1.00	.30	1	.00	2.71
Susceptible to Change	.71	.30	1	.02	2.03
Frame of Reference	1.12	.21	1	.00	3.07
Constant	−2.19	.82	1	.01	.11

Note. −2 Log likelihood = 978.008; model X^2 = 553.390; *df* = 24; *p* < 0.001.

Table B4. Significant predictors of *estar* at level 4

Variable	Coefficient (B)	SE	df	Significance	Estimated odds ratio
Task			2	.00	
Task (1)	1.92	.63	1	.00	6.80
Task (2)	−1.77	.58	1	.00	.17
Accuracy			2	.64	
Acc(1)	4.67	38.91	1	.90	106.77
Acc(2)	10.57	11.27	1	.35	38990.82
Semantic Transparency			4	.03	
Sem(1)	−.11	.94	1	.91	.90
Sem(2)	−1.09	.74	1	.14	.34
Sem(3)	−.85	.53	1	.11	.43
Sem(4)	−2.22	.76	1	.00	.11
Adjective			10	.00	
Adj(1)	1.76	.79	1	.03	5.80
Adj(2)	1.01	.64	1	.12	2.73
Adj(3)	1.01	.70	1	.15	2.75
Adj(4)	.04	1.08	1	.97	1.04
Adj(5)	.70	1.22	1	.56	2.02
Adj(6)	−.27	.60	1	.66	.77
Adj(7)	−1.25	.73	1	.08	.29
Adj(8)	2.67	.91	1	.00	14.42
Adj(9)	−.70	.76	1	.36	.50
Adj(10)	−9.18	40.30	1	.82	.00
Dependence on Experience	1.67	.54	1	.00	5.31
Susceptibility to Change	2.33	.58	1	.00	10.25
Frame of Reference	.93	.33	1	.01	2.53
Constant	−6.08	1.42	1	.00	.00

Note. −2 Log likelihood = 325.595; model X^2 = 257.064; *df* = 21; *p* < 0.001.