

A Longitudinal Study of Copula Choice: Following Development in Variable Structures

Kimberly L. Geeslin and Pedro Guijarro-Fuentes
Indiana University and University of Plymouth

1. Introduction

The second language acquisition of the copula contrast in Spanish (i.e., the distinction between *ser* 'to be' and *estar* 'to be') has often been studied in terms of stages of acquisition. Such stages describe the path that learners take in acquiring all of the various functions of these two copular verbs (Briscoe, 1995; Gunterman, 1992; Ryan & Lafford, 1992; VanPatten, 1985, 1987). In contrast to the analyses of learner accuracy in the context of second language acquisition, sociolinguistic research has demonstrated that in the [copula + adjective] structure there is actually a great deal of variation among native speakers, such that *estar* has come to be used with greater frequency and in a greater number of contexts than was previously acceptable (Cortés-Torres, 2004; Díaz-Campos & Geeslin, 2004; Guijarro-Fuentes & Geeslin, 2006; Gutiérrez, 1992, 1994a, 1994b, 2003; Silva-Corvalán, 1986, 1994). In response to these findings in sociolinguistic research, recent investigations into the second language acquisition of copulas with adjectives have sought to describe copula use by identifying the linguistic and social predictors of the appearance of *estar*, rather than attempting to assess the accuracy of that use (Geeslin, 2000, 2003, 2005; Geeslin & Guijarro-Fuentes, 2006). Additionally, some studies have incorporated native speaker judgments into the design and have distinguished between obligatory and variable contexts (Geeslin, 2001; Geeslin & Guijarro-Fuentes, 2004). This allows researchers to assess accuracy in obligatory contexts while also tracking development in terms of frequency of copula choice in variable contexts. To date, relatively little is known about how learner's use of the copulas changes over time in obligatory and in variable contexts. The current study seeks to explore this issue further. Like other studies, the goal of the current investigation is to learn more about the process of acquisition of the contrast between *ser* and *estar*. What makes the current study unique is that it makes use of native speaker data to distinguish between obligatory and variable contexts and it is a longitudinal study with four different test points over a period of 3 years.

The current article begins with a brief overview of variation in the [copula + adjective] construction, followed by an outline of the key findings of the research conducted so far in which a distinction is made between obligatory and variable contexts in the second language acquisition of Spanish copula choice. The subsequent sections describe the research questions that guided the current investigation, along with a description of the participants, elicitation instruments and methods of analysis. Following the presentation of our findings, we address the implications of our research and directions for future study.

2. Variation in the [copula + adjective] structure

It has long been known that within the [copula + adjective] structure, the use of *estar* has been extended to a wider variety of contexts (see Vañó-Cerdá, 1982 for examples from as early as the 12th century). This language change has been attested in the United States (Silva-Corvalán, 1986), Mexico (Gutiérrez, 1992), Venezuela (De Jong, 1993), and Puerto Rico (Ortiz López, 2000). More recently, researchers have focused on the linguistic and social predictors of the use of *estar* to compare and contrast copula use across populations (Cortés Torres, 2004; Díaz-Campos & Geeslin, 2004; Geeslin, 2003, 2005; Geeslin & Guijarro-Fuentes, 2006). This body of research has shown that even those

varieties that are considered to be conservative, such as those spoken in regions where Spanish is not in contact with other languages, show a tremendous amount of variation both within and across speakers. Falk (1979) proposed that copula choice is actually made according to a speaker's desire to highlight particular aspects of the discourse context. This idea has been incorporated into recent research and it has been shown that speaker choices depend on semantic and pragmatic features of the discourse context, such as whether or not the referent is compared to itself at another point in time, whether the referent is presented in a stage-level or individual-level predicate and whether or not the speaker has immediate experience with the referent. Although the class of the adjective with which the copula is paired remains important, it is no longer believed that the properties of the adjective alone determine the selection of *ser* and *estar* (see Geeslin, 2005 for a complete discussion). Given these developments in the theoretical description of the copulas and the sociolinguistic evidence that supports them, SLA research has had to respond accordingly.

3. The SLA of copula choice: Obligatory and variable contexts

As mentioned in the preceding introduction, early research on the second language acquisition of copula choice focused on the acquisition of several copular functions, using an error analysis to arrive at stages of acquisition (Briscoe, 1995; Gunterman, 1992; Ryan & Lafford, 1992; VanPatten, 1985, 1987). More recently, this focus has shifted from an analysis of accuracy to an analysis of use, in response to recent findings in sociolinguistics (Geeslin, 2000, 2003, 2005; Geeslin & Guijarro-Fuentes, 2006). Studies focusing on an analysis of use generally sought to identify those linguistic and social factors that best predicted copula choice for a group of learners. These predictors could then be compared to other groups of learners at different levels of development or in different learning contexts or to an appropriate native speaker group. Thus, the differences between native and non-native speakers as well as the process of acquisition were described in terms of varying predictors and the varying predictive weights of those factors. While this approach is responsive to the developments in studies of native speaker variation, it still fails to distinguish between those contexts that require *ser*, those that require *estar* and those that allow variation. The benefit in making such a distinction is that it allows for learner data to be evaluated in terms of accuracy without relying on the judgments of a single researcher to determine such accuracy. Moreover, it is possible that contexts where a single copula is obligatory are acquired at a different rate, or with different ease than those contexts that allow variation. Thus, such an approach allows researchers to learn more about the acquisition of copula choice and the acquisition of variation in general.

The first study to make use of the contrast between obligatory and variable contexts was Geeslin (2001). Along with 72 English-speaking learners of Spanish, the study included data from 10 native speakers of Spanish from 5 different countries of origin. The learners belonged to four different proficiency levels ranging from beginner to intermediate. Data were elicited using a 15-item contextualized preference task that asked learners to indicate a preference for either *ser*, *estar* or for both in a given context. Native speaker responses were used to identify obligatory contexts (those where response was unanimous) and variable contexts (those where native speakers did not agree unanimously). It was found that learners at the lowest proficiency level had an accuracy rate of 68 percent in contexts that required *ser*, and this decreased to 52 percent for the highest level. In contrast, contexts that required *estar* had a 45 percent accuracy rate at the lowest proficiency level and a 58 percent accuracy rate at the highest level. This increase in accuracy across time in contexts that required *estar* was shown to be significant. It was concluded that the overgeneralization of *ser* led to accuracy rates that were higher for *ser* at the earlier stages and this explained the apparent decrease in accuracy across time. In contrast, learners gradually work *estar* into the grammar over time, leading to increased rates of accuracy. No results were provided for the contexts where native speakers did not agree.

Following Geeslin (2001), Geeslin & Guijarro-Fuentes (2004) sought to further examine the changes in obligatory contexts over time while at the same time accounting for variable contexts in the analysis. The participants in that study were 7 beginning learners of Spanish whose first language was English and 19 native speakers of Spanish. Data were elicited using a 28-item contextualized

preference task as well as a background questionnaire and a proficiency test. As in Geeslin (2001) the native speaker data were used to identify obligatory and variable contexts. The learners were tested once in the second semester of their first year of study, after about 96 hours of instruction, and again during their third semester of study, after an additional 40 hours of instruction. It was found that the accuracy rates in contexts that required *ser* did not change at all from the first time of testing to the second (both showed 68 percent accuracy). Likewise, in contexts that required *estar*, accuracy rates changed from 74.3 percent at the first time to 80 percent at the second time, but this difference was not significant. The greatest change over time was found in the variable contexts where use of *estar* increased significantly from 57 percent to 76.5 percent. Thus, it was concluded that the variable contexts warranted further study because that is where the greatest changes were taking place. It was further hypothesized that the change in accuracy of *estar* would reach significance over a greater period of time. Finally, it was suggested that research over a greater period of time would be necessary in order to determine when learners began to improve in *ser*-required contexts, whether the increase in use of *estar* remained steady over time and whether or not a u-shaped pattern of development might be found in either the obligatory or the variable contexts. The current study was designed to examine these very questions by collecting data from these same learners at later points in time and comparing results from these first two test points to results from two additional testing times.

4. The current study

As stated previously, the current study was designed to examine the second language acquisition of copula choice in [copula + adjective] contexts across a three-year time period while at the same time accounting for the distinction between obligatory and variable contexts. Taking into account our previous analyses of these data, the current study was guided by the following research questions:

1. Are there significant changes in contexts where *ser* is required after more time of exposure to Spanish has passed?
2. Are there significant changes in contexts where *estar* is required after more time of exposure to Spanish has passed? (as in Geeslin, 2001)
3. Do learners continue to increase their use of *estar* in variable contexts across time?

4.1. Participants

In the current study, two groups of participants were included: a native Spanish-speaking group, and a native English-speaking group of second language learners of Spanish. The group of 19 native speakers of Spanish was included so that obligatory contexts, those where native speakers unanimously selected either *ser* or *estar* could be distinguished from variable contexts in which native speaker responses varied across participants. The native speaker group was residing in Granada at the time of data collection, but their cities of origin included Granada, Jaén, Madrid, Murcia, and León. None of these regions is associated with bilingualism with a regional language. All but 6 of these participants were female, and the age of participants ranged from 20 to 46 (mean=32.47, s.d. = 6.60). All but 4 participants had received higher education, and all had completed secondary education. Some speakers in this group also had experience with Catalan, French, German, Portuguese and English. Participants were recruited through academic and social networks and participated voluntarily. All data collection with this group took place in Granada, Spain.¹

The second group was composed of 7 English-speaking learners of Spanish who started studying Spanish as part of their degree program. The chronological age of participants ranged from 20 to 47 (mean = 28.7; s.d. =11.02), and none had any previous knowledge of Spanish at the start of the project.² Four of the 7 participants were female. During the first two years of instruction (the first 3 testing times), all learners had four hours of formal classes per week during two 12-week semesters, equal to approximately 96 hours of formal instruction per academic year and data collection took place in the UK. During the third year of instruction (test time #4) 4 of the 7 participants spent 4 months in Spain while the remaining 3 continued their studies in the UK, still receiving the 4 hours of formal

instruction per week. Aside from experience abroad, all participants had the same native-speaking instructors during the entire length of the study. Additional details about the 4 testing times are provided in the following section.

4.2. Elicitation instruments

All participants in the current study completed the same tasks, with the exception of the proficiency assessment task which was completed only by the second language group. All tasks were completed in the presence of one of the researchers and were completed in one hour or less. Participants began by completing a background questionnaire that asked questions regarding language use, language study and individual characteristics. There were slight differences in the versions presented to each group. For example, only the non-native group was asked to give information as to when they began studying Spanish. Additionally, while the native-speaking group completed the instrument in Spanish, the learners completed the instrument in English.

Next, the second language group completed an assessment of proficiency containing 43 items focused on basic aspects of Spanish grammar (e.g., the preterit/imperfect distinction, the use of the subjunctive, etc.). The task was a multiple-choice contextualized task completed in Spanish, requiring participants to select the option (preposition, verb form, etc.) that best filled the blank in the sentence provided in the test item. Participants' scores ranged from 24 to 42 and increased significantly ($p < .001$) across time. The complete results for this test for each participant across time are described in Table 1.

Table 1. Results of proficiency test for each participant across time

Time	Part 1	Part 2	Part 3	Part 4	Part 5	Part 6	Part 7	Mean	s.d.
1	26	29	30	27	34	33	30	29.86	2.702
2	39	33	32	33	37	24	30	32.57	4.511
3	42	39	39	36	42	37	40	39.29	2.2124
4	41	40	33	31	39	40	38	37.43	3.589

Note: Part = participant. $X^2=1381.33$, $df=45$, $p < .001$, Cramer's $V = .766$.

The final written task used in the current study to gather data was a contextualized preference task. The task included 28 items. Each item contained a paragraph-length context which established the discourse situation and manipulated several of the contextual features that have been found to be associated with the use of *estar* (see Geeslin, 2005 for details).³ The entire instrument took the form of a story about two college roommates. Following each paragraph, one of the characters in the story asked a question and the other character responded. This response was given in the form of two sentences that were identical except for the copula that each contains. The participant was directed to indicate if he or she preferred sentence A, sentence B or if both were acceptable. The order of presentation of each copula was randomly varied throughout the instrument. All participants completed all 28 items. See Figure 1 for a sample item (Shown in translation).

Figure 1: Sample item from the contextualized preference task.

1. Paula and Raúl are going to a restaurant tonight. Paula is talking from her bedroom while she gets dressed and making plans with Raúl, who is in the living room. When she comes out of the bedroom she asks Raúl:	
Paula: Would you like to go in my car?	
A. Raúl: Wow! How pretty you are (<i>estar</i>)!	___ I prefer sentence A
B. Raúl: Wow! How pretty you are (<i>ser</i>)!	___ I prefer sentence B.
	___ I prefer A and B.

The native speaking group completed the tasks only once whereas the non-native group completed the tasks at 4 different times. The first two collection times, described above in Geeslin & Guijarro-Fuentes (2004) took place during the second semester of the first year of instruction (after about 96 hours of instruction) and again during the first semester of the second year (after about 136 hours of instruction). In addition, we collected data during the second semester of the second year (after 186 hours of instruction) and finally during the third year of instruction (after 6 additional months of instruction). At the last time of testing 4 of the 7 participants had spent 4 months in Spain while the others had continued their studies of Spanish in England.

4.3. Analysis

To answer the research questions posed previously, the data were analyzed in several ways. First, the overall distribution of copula choice across time was tabulated. To allow further statistical analysis it was then necessary to recode the dependent variable to include contexts where participants accepted *estar* and those where they did not. This binary variable alleviates the difficulties created by the small number of tokens of the ‘both’ response. Next, the native speaker data were used to divide the items according to whether responses were unanimous (i.e., obligatory contexts) or variable.⁴ The obligatory contexts were further divided between those that required *ser* and those that required *estar*. In all three of those contexts chi-square (X^2) tests were conducted to determine whether or not the change across time was significant.⁵ Finally, an individual item analysis was conducted for each of the three context types in order to determine whether the overall patterns of development adequately represented the patterns exhibited on each item or whether further variation was present.

5. Results

A total of 196 tokens were collected at each of the four times data were elicited yielding a total of 884 tokens. Of these tokens, only 16 were instances in which participants responded with the ‘both’ option. Consequently, these tokens were recoded into a dependent variable that distinguished between responses where *estar* was accepted and responses where it was not (see Geeslin, 2005 for the rationale for this approach). These categories have been called ‘*estar* allowed’ and ‘*estar* not allowed’ in previous research and that terminology will be maintained here. The distribution of these responses across time is described in Table 2.

Table 2. Response Type (*Estar* Allowed vs. Not allowed) across Time

Response	Time 1		Time 2		Time 3		Time 4	
	#	%	#	%	#	%	#	%
<i>Estar</i> allowed	102	52	123	62.8	127	64.8	109	55.6
<i>Estar</i> not allowed	94	48	73	37.2	69	35.2	87	44.4
Total	196	100	196	100	196	100	196	100

Note. $X^2=8.693$, $df=3$, $p = .034$, Cramer’s $V = .105$.

The distribution of these tokens is significantly different across time, but neither use of *ser* nor of *estar* can be described as a linear increase or decrease. Instead, it can be seen that the use of *ser* begins around 50 percent and then decreases during the next two phases, increasing only at the fourth time of testing but not to the level of the first testing time. In contrast, the use of *estar* shows the opposite pattern. Use of *estar* also begins around 50 percent but gradually increases toward the second and third testing times, decreasing only at the fourth testing time but staying above the starting level. Thus, the use of *ser* over time can be described as u-shaped while the use of *estar* is the opposite (an inverted u). This result is not unlike what has been seen for the development of other grammatical structures (Ellis, 1994, p.303), but has yet to be shown for copula choice.

Once the overall distribution of responses was established, the native speaker responses were divided to distinguish between obligatory and variable contexts. An obligatory context was defined as any item where all 19 native-speaking participants provided the same response. A total of 14 of the 28

contexts were found to be variable, 5 were found to require *estar* and 9 were found to require *ser*.⁶ The distribution of responses in each of these contexts was examined individually. Table 3 describes the learner responses in contexts where *ser* was the unanimous response among native speakers.

Table 3. Response Type (*Estar* Allowed vs. Not allowed) across Time in Contexts that Require *ser*

Response	Time 1		Time 2		Time 3		Time 4	
	#	%	#	%	#	%	#	%
<i>Estar</i> allowed	20	31.7	20	31.7	19	30.2	13	20.6
<i>Estar</i> not allowed	43	68.3	43	68.3	44	69.8	50	79.4
Total	63	100	63	100	63	100	63	100

Note. 9 contexts require *ser*. $X^2 = 2.644$, $df = 3$, $p = .45$, Cramer's $V = .102$.

Table 3 illustrates that the *estar* responses in contexts where *ser* is required decreased gradually over time, but that this decrease largely takes place between the third and fourth testing times. Thus, only at the last time of testing do we see much improvement in learner accuracy on these items. This increase across time was not found to be significant. What is most interesting about these results is that learners are still relatively inaccurate on these items. Previous research showed that *ser* was overgeneralized in early stages to the extent that accuracy rates may have appeared higher than they actually were, and only Briscoe (1995) claimed there were functions of *ser* acquired after functions of *estar*. The current study shows that learners still have a long way to go in learning *ser*, even after three years of study of the language. In order to further examine these contexts an item analysis was conducted and the responses across time on items that require *ser* are described in Table 4.

Table 4. Percent Correct across Time in Contexts that Require *Ser*

Time	# 6	# 7	# 8	# 10	# 13	# 17	# 20	# 21	# 25
1	57	86	71	43	86	71	100	57	43
2	100	86	29	71	71	71	86	71	29
3	71	86	14	86	43	71	86	86	86
4	86	57	57	100	100	71	100	86	57

Table 4 illustrates that there are many different patterns of development within this group. While items 10, 17 and 21 show steady use or improvement across time, items 7 and 8 show lower accuracy at the last phase than at the first. Items such as 6, 13, and 25 show sharp changes (either toward accuracy or away from it) at various points in the process of development. Items 20, 13, and 8 show the u-shaped development exhibited in the overall copula use across all items, regardless of context. Because there were only 7 participants, some apparent changes may not be as sharp as they appear in Table 3, but it is important to note the high degree of variability found among these items.

Table 5 illustrates the distribution of copula choice across time in contexts where *estar* was required. Like the use of *ser* in *ser*-required contexts, the use of *estar* in *estar*-required contexts does increase over time. In contrast to the *ser*-required data, the largest jump takes place between the second and third testing times, rather than the third and fourth. This change across time approaches significance but does not quite reach the .05 level. Another important result for these contexts is that learners do reach a relatively high level of accuracy. It will be recalled that previous studies on the stages of acquisition used 90 percent as the cut-off point for indicating that a structure had been acquired (VanPatten, 1987). Thus, the learners in this study improve sooner on their use of *estar* in *estar*-required contexts than they do with *ser* in *ser*-required contexts and they reach a relatively high level of accuracy on this structure by the third and fourth test times.

Table 5. Response Type (*Estar* Allowed vs. Not allowed) across Time in Contexts that Require *Estar*

Response	Time 1		Time 2		Time 3		Time 4	
	#	%	#	%	#	%	#	%
<i>Estar</i> allowed	26	74.3	28	80	32	91.4	33	94.3
<i>Estar</i> not allowed	9	25.7	7	20	3	8.6	2	5.7
Total	35	100	35	100	35	100	35	100

Note. 5 contexts require *estar*. $X^2 = 7.339$, $df = 3$, $p = .062$, Cramer's $V = .229$.

As with the data for *ser*-required contexts, an item analysis was conducted for contexts that require *estar*. The percentage of use of *estar* in these contexts is described in Table 6.

Table 6. Percent Correct across Time in Contexts that Require *Estar*

Time	Item 2	Item 4	Item 12	Item 18	Item 22
1	100	43	57	100	71
2	100	71	57	100	71
3	86	86	86	100	100
4	100	100	100	100	71

Table 6 shows that in general there is improvement over time on items that require *estar*. While there is a jump in accuracy on item 22 at time 3 and a dip in accuracy at time 3 on item 2, these jumps do not change the overall improvement across time. What is most remarkable about these results is that learners reach such a high level of accuracy on these items by the end of the study. All but 1 of the items shows perfect accuracy at the highest level of testing. In looking at the responses across time, it is also likely that the failure of the change across time to reach significance is due to the relatively high level of accuracy demonstrated on items 2 and 18 at all four testing times.

In addition to the items where accuracy was possible to assess, we analyzed those contexts where native speaker responses varied and no single correct answer could be identified. The responses to these variable contexts are summarized in Table 7.

Table 7. Response Type (*Estar* Allowed vs. Not allowed) across Time in Variable Contexts

Response	Time 1		Time 2		Time 3		Time 4	
	#	%	#	%	#	%	#	%
<i>Estar</i> allowed	56	57	75	76.5	76	77.6	63	64.3
<i>Estar</i> not allowed	42	42.9	23	23.5	22	22.4	35	35.7
Total	98	100	98	100	98	100	98	100

Note. 14 contexts are variable. $X^2 = 13.76$, $df = 3$, $p = .004$, Cramer's $V = .185$.

Table 7 indicates that the use of *estar* changes significantly over time. Like the overall distribution of responses for all items, these items show an increase in *estar* toward the second and third testing times followed by a decrease in selection of *estar* at the last time of testing. Thus, the pattern for selection of *estar* in variable contexts is a u-shaped pattern, but one in which the last test time does not dip as low as the rate of selection at the first test time. An analysis of each of the items in the group was conducted according to the frequency with which native speakers selected *estar* in the same contexts. Thus, the item analysis includes items where native speaker use of *estar* is high (above 80 percent), where it falls in the middle range (between 60 and 80 percent) and where native speaker selection of *estar* is low (below 33 percent). The results for the first of these contexts are summarized in Table 8. Although it appears that native speakers selected *estar* 100 percent of the time on three of the items in this group, which would make the item obligatory rather than variable, this is because the results reflect the binary dependent variable where the both response was collapsed with the *estar* response. Thus, at least one native speaker responded that both options were acceptable in these contexts.

Table 8. Percent '*Estar Allowed*' Where NS Use Ranges from 80 to 100 Percent

Time	Item 1	Item 16	Item 23	Item 27	Item 28
1	71	86	43	29	29
2	86	100	71	57	71
3	86	100	86	86	57
4	86	86	57	57	43
NS	90	100*	100*	89.5	100*

*NS only show 100% when 'both' response is combined with *estar* (i.e., they are variable).

Table 8 shows that within variable contexts where native speaker selection of *estar* is relatively high, learner selection of *estar* tends to be lower than that of native speakers. On items 1 and 16 learners do approach native speaker rates of selection, but on items 23, 27 and 28 learners show a spike in use around the second or third testing times and a subsequent dip in selection of *estar* at the last time of testing. The end result is a rate of selection well below that shown for native speakers. Thus, the task in these variable contexts for language learners is to continue to increase the rate with which they select *estar*.

Table 9 shows the distribution of responses in variable contexts where native speaker selection of *estar* falls between 60 and 80 percent. It can be seen that on these items learner rates of selection of *estar* are similar to those of native speakers. In cases such as items 3, 15 and 19, learners first overshoot the target and select *estar* more often than the natives, but at the time of the fourth testing are back down to similar rates of selection. Only item 5 shows a u-shaped pattern of development, with a dip in selection at the second time of testing followed by a subsequent increase at times three and four. In comparison to the variable contexts on which native speaker rates of selection are quite high, contexts where native speaker rates of selection of *estar* are in the mid-range show a closer approximation to native speaker norms.

Table 9. Percent '*Estar Allowed*' where NS Use Ranges from 60 to 80 Percent

Time	Item 3	Item 5	Item 15	Item 19
1	29	57	57	86
2	100	14	100	86
3	100	71	86	86
4	86	71	86	71
NS	79	74	68	79

Finally, Table 10 shows the rates of selection of *estar* in variable contexts where native speaker use is relatively low. It will be noted that while the mid-range rates of selection begin at 60 percent, the highest value in the low range category is 33 percent. This is because no tokens exist for which native speaker rates of selection were between these two values.

Table 10. Percent '*Estar Allowed*' Where NS Use is below 33 Percent

Time	Item 9	Item 11	Item 14	Item 24	Item 26
1	29	71	86	57	71
2	29	86	86	86	100
3	29	71	86	100	43
4	14	71	86	57	43
NS	5.3	16	32	16	5.3

Table 10 demonstrates that non-native speaker use of *estar* in variable contexts where native speaker rates of selection of *estar* are low tends to be inappropriately high. Although item 9 shows a steady decrease over time toward a lower rate of selection of *estar*, the other four items show non-native rates of selection that are quite different from the native speaker rates. Item 14 shows no change at all across time and item 24 shows a u-shaped pattern such that learners were no closer to the native speaker norm at test time four than at the first test time. Items 11 and 26 show initial increases in the

use of *estar* and then subsequent decreases in use which move toward the norm but do not approximate it very closely. Because these are indeed variable contexts, it would be inappropriate to expect that learner rates of selection matched the native speaker rates of selection exactly. In fact, because native speakers acknowledge that both responses are possible, any response given by the learners is within the range of permissible responses. Still, it can be concluded that as a group, there are still differences between the native and non-native speakers on most items.

The last variable that was examined in the current study was the difference between those participants who studied abroad during the third year of study and those that did not. About half of the participants (4 out of 7) participated in a study abroad program in Granada, Spain. This group completed the test instrument during that time abroad whereas the other participants completed the final test instrument in the UK. The differences in responses between these two groups are shown in Table 11.

Table 11. Response Type at Time 4 by Study Abroad

	Study abroad		No study abroad	
	#	%	#	%
Estar allowed	52	46.4	57	67.9
Estar not allowed	60	53.6	27	32.1
Total	112	100	84	100

Note. $N=196$, $X^2 = 8.929$, $df = 1$, Cramer's $V = .213$, $p = .003$

Table 11 demonstrates that the group that participated in the study abroad program showed a considerably lower rate of selection of *estar* than the group that did not. This difference was shown to be significant using a X^2 test. What this means in terms of overall rates of selection across time is that the group who studied abroad shows a much sharper curve between the third and fourth test time than the group that did not. The study abroad group shows a sharper increase in the use of *ser* at the fourth test time and a sharper decrease in the use of *estar* than the group that did not study abroad. The implications of these results and their connections to the research questions that guided the current study will be assessed in the following section.

6. Discussion

To summarize, the current study found a u-shaped pattern of development for copula choice when looking at overall rates of selection of the copulas and when looking at variable contexts. This curve was sharper for the group who studied abroad. In addition, it was found that rates of accuracy for *ser* in obligatory contexts start lower than rates of accuracy for *estar* (68% vs. 74%), and the ultimate accuracy for *ser* at the fourth test time was lower than it was for *estar* (80% vs. 94%). Looking more closely at individual items that require *ser*, there is evidence of a gradual increase in accuracy on some items as well as u-shaped patterns that begin with overgeneralization, continue with a dip in rates of selection and ultimately show an increase at the later two testing times. Still, 2 of the 9 items show a lower rate of accuracy at the fourth test time than at the first. The item analysis of contexts that require *estar* showed a general increase in rates of selection of *estar* over time, with all but 1 of the 5 items reaching 100 percent accuracy by the last testing time. Although improvement across time was not shown to be significant this is likely due to the fact that 2 of the 5 items show a high rate of accuracy as early as the first test time. Finally, in examining variable contexts, learners show varying degrees of correspondence with native speaker rates of selection of *estar* depending on whether native speaker use is high, mid-range or low. On contexts where native speaker rates of selection of *estar* are relatively high, learners show an increase toward the norm, but are farther from the native speaker rates of use on 4 of the 5 items at the last time of testing than they were at least one earlier test time. On contexts where the native speaker rates of selection of *estar* were between 60 and 80 percent, non-native speakers show a decrease in use of *estar* toward the norm and reach a relatively similar rate to that of the native speakers by the last test time. Lastly, in contexts where native speaker rates of selection of *estar* are relatively low, learners tend to overshoot the target and although they show an

eventual decrease in use toward the norm, learner rates of selection of *estar* are farthest from the rates of selection of native speakers on these items.

It will be recalled that the current study was guided by three research questions. The first sought to determine whether or not there were significant changes in rates of use in contexts where *ser* is required after more time of exposure to Spanish had passed. Previous research such as Geeslin (2001) and Geeslin & Guijarro-Fuentes (2004) had shown that in early stages of development the overgeneralization of *ser* (i.e., initially high rates of use of *ser*) tends to obscure any development that might be taking place. It was proposed that a significant result might never be obtained because this overgeneralization would eventually lead to correct use and inaccuracy rates would never be terribly high. In fact, it was the case in the current study that there was not a significant change in rates of selection of *ser* across time in contexts that require *ser*. The surprising part of this result is that the accuracy rates for *ser* were quite low. In fact, the average accuracy rate on all 9 items in this context after 3 full years of study was still well below the 90 percent that would indicate that the structure had been acquired. Thus, the idea that *ser* is acquired early is questionable. The only previous study to have shown that any function of *ser* was acquired after any function of *estar* began to be acquired was Briscoe (1995) who noted that the use of *ser* in locatives (for events) and in the passive voice tended to be acquired late. Our study is the first to show that this is also true for adjectival contexts where *ser* is required.

The second research question that the current study was designed to answer was whether or not the rates of selection in contexts that require *estar* changed significantly across time. Geeslin (2001) found that changes in such contexts were significant and Geeslin & Guijarro-Fuentes (2004) did not, but suggested that this might be due to the relatively low proficiency level of the learners at the two testing times in that study. Thus, it was hypothesized that over a longer period of time significant changes in the rates of use of *estar* would be observable. In fact, the current study found that while the rates of change across time approached significance in the 5 contexts where *estar* was required, the statistical test did not yield a significant result. Nevertheless, learners did improve across time and reached 100 percent accuracy on all but 4 of the 5 items in this group. It is likely that this result was not statistically significant because learners began with a very high rate of accuracy on 2 of these items. Thus, the high rates of accuracy in contrast with the lower rates for *ser* further demonstrate that the notion that *ser* is acquired prior to *estar* is erroneous. Instead, a much closer analysis of the contexts included in each of these environments is required.

The final research question that the current study sought to address was whether or not learner selection of *estar* continued to increase over time in variable contexts. Results from Geeslin & Guijarro-Fuentes (2004) showed a significant increase in the rates of selection of *estar* in these contexts and it was hypothesized that this use would continue to increase over time. In fact, the changes across time in these contexts were shown to be significant, but the rates of use are better described as u-shaped rather than a steady linear increase across time. At the fourth time of testing there was a dip in the rates of selection of *estar*, similar to the rates of use overall. When looking at these contexts according to rates of native speaker use, it was seen that learners least approximated the native speaker rates in contexts where native speaker selection of *estar* was lowest. In relating this result to the acquisition of variation in general, it can be hypothesized that learners are able to identify a context as variable (because they know *estar* is allowed in these contexts) before they are sensitive to the rate at which such a variant is actually supplied (because their rates of selection of *estar* tend to overshoot those of the native speakers). It would be interesting to test this hypothesis with other structures to see whether or not the same pattern emerges.

In addition to answering the research questions posed in the current study, this investigation has also made a methodological contribution to the work on copula choice. The incorporation of a native speaker population to identify obligatory contexts is not one that should be taken lightly. These contexts are comprised of a complex group of factors each of which has a degree of influence on the copula selection. Thus, no single linguistic feature is present in all of the contexts that require a single copula (see Geeslin 2001 for evidence of this claim). Instead, only native speaker performance on the same task that learners complete is sufficient for arriving at this distinction. In addition to using a native speaker baseline to determine which contexts were obligatory, the native speaker rates of use also made it possible to examine variable contexts in greater detail than has previously been done.

Finally, the longitudinal nature of our study has made it possible to track learners across 3 years of study, including a study abroad experience, and this makes it possible to talk about development of copula choice in all three contexts over time.

7. Conclusions and future directions

The current study has demonstrated that the apparent patterns of u-shaped development for both *ser* and *estar*, shown in the overall rates of selection and in the variable contexts, disappear when contexts are examined in smaller groups according to patterns of native speaker use. Both *ser* and *estar* show increases in accuracy across time. While neither shows significant improvement across time, accuracy rates for *estar* are actually quite high. This study is one of the first to document the fact that not all contexts for *ser* are easily acquired and should lead researchers to rethink assumptions about how *ser* is acquired across time. In variable contexts, learners approximate native speaker rates of use fairly well when such use falls between 60 and 80 percent but tend to select *estar* with a much higher frequency than native speakers in contexts where native speaker use is low. Future studies should examine copula choice in variable contexts based on comparable native speaker rates of use in order to ensure the appropriate level of detail of analysis.

There is a great deal of research left to be done in this area. First, we note that the current participant group was chosen specifically because of their homogeneity, but provides a relatively small sample size. In larger studies, it would be interesting to explore additional individual variables. A larger participant sample would make it possible to include individual variables and to submit these data to statistical tests such as the regression analysis. The current study served to formulate hypotheses about the acquisition of this variable structure and it is important to determine whether or not such research extends to other variable structures. Thus, future research should replicate the current study with a different structure as the focus of analysis. Finally, research on copula choice in Spanish remains largely limited to English-speaking learners of Spanish and the current study is not different. We still know very little about what the process of development for learners with different first language backgrounds might look like (see Geeslin & Guijarro-Fuentes, 2005a for an example). Future research should include participants whose first languages differ from both English and Spanish in order to broaden our general understanding of the acquisition of this structure. In sum, the current study has both contributed new findings on the second language acquisition of copula choice and of variable structures in general while at the same time generating new questions for future research.

Notes

*We would like to thank the participants in 2005 Joint Conference on Hispanic Linguistics and Acquisition (The Pennsylvania State University), the anonymous reviewers and the volume editors, and also the British Academy for the research grant which allowed us to conduct this research.

1. It was suggested that other characteristics, such as level of education or gender, of this native-speaking group might also contribute to variation among native speakers, and that this might be problematic given the relatively small sample size and an uneven distribution of individual characteristics within the group. In fact, a previous analysis of these 19 speakers showed that no social variables (e.g., level of education, gender, age) were significant predictors of the use of *estar*. Moreover, the 19 native-speakers included in the current study are part of a larger database to which an additional 63 native speakers of Spanish living in monolingual regions of Spain have been added since our original analysis was conducted (N=82). Analyses of the individual participant characteristics in this group have shown that only gender is a significant predictor of the selection of *estar*, such that women select *estar* more often (Geeslin & Guijarro-Fuentes, 2005b). Because the current study does not seek to examine the predictors of *estar*, but rather the range of variation that exists on a given test item (each of which contains a context that specifies the categories of the variables that have been found to predict the selection of *estar*), we conducted an additional analysis to ensure that the 19 participants included here were representative of the group as a whole, and that having included a more homogeneous group would not have changed the results. On the 6 *estar*-required items, adding

the remaining 63 participants to the analysis only produced one response on each of 2 items that did not agree with the group. This response was given by the same participant. Thus, unanimity was maintained among the other 81 participants. Given that the use of *estar* is being extended, this result is expected. On the 9 *ser*-required items, three remained unanimous, 3 showed a single participant who disagreed with the other 81 and the remaining 3 showed no more than 5 participants who disagreed with the rest of the group. Although these results exhibited slightly more variation than the smaller sample, there are several points worth making. First, given that the original population was one-fourth the size of the larger group, the amount of additional variation found is quite small. Secondly, the disparate responses on the *ser*-required items come from both men and women and from speakers with only secondary education and with higher education. Thus, the identification of obligatory contexts is not linked to individual characteristics of this type.

2. Although the number of participants is relatively small, this group was chosen for their homogeneity in language learning experience and social characteristics, and because this group followed the same path of study over the course of three years.

3. Categories of variables such as frame of reference, susceptibility to change and adjective class were evenly distributed across the instrument. An analysis of this distribution shows that no single syntactic, semantic or pragmatic feature can predict the classification of an item as obligatory or variable. In other words, the categories of these variables are distributed across both obligatory contexts and variable contexts.

4. It should be noted that in collapsing the dependent variable, individual variation (speakers who select 'both') is being grouped with variation across individuals (speakers who disagree in their selection of a single copula). This is not problematic, however, because the purpose of the classification of the instrument is to identify contexts that allow variation, regardless of the type.

5. Although the data in the current study have been tabulated and presented to the reader in the form of percentages for the sake of clarity, this is not the form in which the data were analyzed. Rather than comparing the changes in the percentage of selection of *estar* for each participant under different conditions (e.g., at different times, with or without study abroad, etc.), each token is included in the analysis and the dependent variable is the copula that was selected. Thus, all of the tokens collected at time one (N=196), for example, are treated as a single group and compared to the tokens collected at times two, three and four. The X^2 test is capable of effectively assessing whether or not the categories of the dependent variable (copula selected) are associated significantly with any of the categories of the independent variable (e.g., time of test).

6. One reviewer was surprised by the amount of variation found among speakers of a single dialect and questions whether or not this is a result of confusion caused by the test items. In fact, the variation is the result of speakers choosing to highlight different aspects of the discourse context. In Figure 1 in the current paper, the example did not prompt unanimous responses. Some native speakers chose to highlight the individual frame of reference presented in this context whereas others chose to treat physical appearance as a more permanent characteristic. Thus, when the discourse features that prompt *estar* are in conflict, native speakers do not make unanimous decisions about how to respond. This result has been quite consistent across several different native-speaking populations. A similar version of this instrument is published in its entirety in Geeslin (2005).

References

- Briscoe, Gregory. (1995). *The acquisition of ser and estar by non-native speakers of Spanish*. Ph.D. dissertation. University of Pennsylvania.
- Cortés-Torres, Mayra. (2004). ¿Ser o estar? La variación lingüística y social de estar más adjetivo en el español de Cuernavaca, Mexico. *Hispania*, 87, 4, 788-795.
- De Jonge, Bob. (1993). (Dis)continuity in language change: Ser and estar + age in Latin American Spanish. In Dijkonigen, F. & Hengeveld, K. (Eds.), *Linguistics in the Netherlands*. (pp. 69-80). Amsterdam: John Benjamins.
- Díaz-Campos, Manuel. & Geeslin, Kimberly. (2004). Copula use in the Spanish of Venezuela. Paper Presented at NWAV, Ann Arbor, MI.
- Ellis, Rod. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

- Falk, Johan. (1979). *Ser y estar con atributos adjetivales*. Uppsala: Alqvist and Wiksell.
- Geeslin, Kimberly. (2000). A new approach to the second language acquisition of copula choice in Spanish. In Ron Leow & Cristina Sanz (Eds.), *Spanish Applied Linguistics at the Turn of the Millennium: Papers from the 1999 Conference on the L1 & L2 Acquisition of Spanish and Portuguese*, (pp. 50-66). Somerville, MA: Cascadia Press.
- _____. (2001). Changing norms, moving targets and the SLA of copula choice. *Spanish Applied Linguistics*, 5 (1-2), 29-55.
- _____. (2003). A comparison of copula choice in advanced and native Spanish. *Language Learning*, 53, 4, 703-764.
- _____. (2005). Crossing Disciplinary Boundaries to Improve the Analysis of Second Language Data: A Study of Copula Choice with Adjectives in Spanish. Munich: LINCOM Publishers.
- Geeslin, Kimberly & Guijarro-Fuentes, Pedro. (2004). Estudio longitudinal del *Ser y Estar* en el español como L2. *Porta Linguarum*, 2, 93-110.
- _____. (2005a). The acquisition of copula choice in instructed Spanish: The role of individual characteristics. In Eddington, D. (Ed.), *Studies in the Acquisition of the Hispanic Languages: Papers from the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, pp. 66-77. Somerville, MA: Cascadia Press.
- _____. (2005b). The SLA of variable structures: Analyzing Obligatory and Variable Contexts. Paper presented at AILA (International Association of Applied Linguistics), Madison, WI.
- _____. (2006). The Second Language Acquisition of Variable Structures in Spanish by Portuguese Speakers. *Language Learning* 56 (1), 53-107.
- Guijarro-Fuentes, Pedro. & Geeslin, Kimberly. (2006). Copula choice in the Spanish of Galicia: The effects of bilingualism on language use. *Spanish in Context*, 3(1).
- Gunterman, Gail. (1992). An analysis of interlanguage development over time: Part II, Ser and estar. *Hispania* 75, 1294-1303.
- Gutiérrez, Manuel. (1992). The extension of estar: A linguistic change in progress in the Spanish of Morelia, Mexico. *Hispanic Linguistics*, 5(1-2), 109-41.
- _____. (1994a). Simplification, transfer and convergence. *Bilingual Review*, 19(2), 111-121.
- _____. (1994b). La influencia de "los de abajo" en tres procesos de cambio lingüístico en el español de Morelia, Michoacán. *Language problems and language planning*, 18, 3, 257-269.
- _____. (2003). Simplification and innovation in US Spanish. *Multilingua* 22, 169-184.
- Ortiz López, Luis. (2000). La extensión de estar en contextos de ser en el español de Puerto Rico: ¿Evolución interna o contacto de lenguas? *BAPLE*. 98-118.
- Ryan, John & Lafford, Barbara. (1992). The acquisition of lexical meaning in a study abroad environment: *Ser + estar* and the Granada experience. *Hispania*, 75, 714-22.
- Silva-Corvalán, Carmen. (1986). Bilingualism and language change: The extension of estar in Los Angeles Spanish. *Language* 62, 587-608
- _____. (1994). *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
- Vañó-Cerdá, Antonio. 1982. *Ser y Estar + Adjetivos*. Tübingen: Narr.
- VanPatten, Bill. (1985). The acquisition of ser and estar in adult second language learners: A preliminary investigation of transitional stages of competence. *Hispania*, 68, 399-406.
- _____. (1987). The acquisition of ser and estar: Accounting for developmental patterns. In Bill VanPatten, Trisha Dvorak, & James Lee (Eds.), *Foreign language learning: A research perspective* (pp. 61-75). New York: Newbury House.

Selected Proceedings of the 9th Hispanic Linguistics Symposium

edited by Nuria Sagarra
and Almeida Jacqueline Toribio

Cascadilla Proceedings Project Somerville, MA 2006

Copyright information

Selected Proceedings of the 9th Hispanic Linguistics Symposium
© 2006 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 1-57473-413-X library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Geeslin, Kimberly L. and Pedro Guijarro-Fuentes. 2006. A Longitudinal Study of Copula Choice: Following Development in Variable Structures. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 144-156. Somerville, MA: Cascadilla Proceedings Project.

or:

Geeslin, Kimberly L. and Pedro Guijarro-Fuentes. 2006. A Longitudinal Study of Copula Choice: Following Development in Variable Structures. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 144-156. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1374.