Comparing Interview and Written Elicitation Tasks in Native and Non-native Data: Do Speakers Do What We Think They Do?

Kimberly L. Geeslin and Aarnes Gudmestad
Indiana University and Virginia Polytechnic Institute and State University

Research on the second language acquisition (SLA) of Spanish has identified grammatical structures for which an analysis of errors for second-language (L2) learners is inappropriate (Geeslin, 2003; Geeslin & Guijarro-Fuentes, 2006; Gudmestad, 2006). This is because the norms of use for such structures are changing, and prescriptive grammars do not coincide with actual language use. Thus, in order to examine such sociolinguistically-variable grammatical features in learner language, researchers have shifted to an analysis of the predictors of use of a given variant, rather than an assessment of accuracy (Geeslin, 2000). Investigations following this approach on structures such as copula choice and mood choice have been largely based on written contextualized tasks (WCT), where use is contextualized and participants indicate a preference for one of the two possible variants. The advantage of this type of task is twofold. First, in comparison with grammaticality judgment tasks, participants are not forced to select one (presumably the only) grammatical sentence from the options provided. Consequently, the WCT is more in line with the idea that variation is indeed an acceptable, and even irrefutable, part of native-like speech. Secondly, in comparison with tasks that elicit less directed production, the WCT assures that each participant will respond to the same tokens (both lexically and in terms of the contextual features that predict selection of a given variant) and that each participant will provide sufficient data for analysis. Nevertheless, while the aforementioned advantages of this methodology are clear (Chaudron, 2003), the degree to which such data generalize to those collected using other elicitation tasks is less widely understood. In other words, what we think speakers do as a result of these studies may not be an accurate view of what speakers do in general or in other elicitation contexts.

The current paper contributes to the relatively extensive body of research in which variation across tasks among L2 learners has been documented (Bayley & Preston, 1996; Bygate, 1999; Ellis, 1987) by exploring cases of sociolinguistically-variable structures, such as use of the Spanish indicative and subjunctive moods and the copulas. Specifically, we compare both the frequency and the predictors of choice[1] of these two structures across tasks for advanced non-native and native-speaking groups. In so doing, our goal is to learn more about the SLA of variable structures in general, about the ways native speaker (NS) and non-native speaker (NNS) variation differ from each other, and about the influence of different elicitation tasks on language.

1. Background
1.1 Task Variation in Second Language

There is a long tradition of investigating variation across tasks in SLA research. Following developments in sociolinguistic theory (see Labov, 1994) interest in variation in L2 data began as an examination of varying attention paid to speech (Tarone, 1988). Dickerson and Dickerson (1977) found systematic differences in the production of English (r) and (z) by Japanese speakers in data elicited through free speech, dialogue reading, and word list reading, and similar effects were found for numerous variants among speakers from many language backgrounds (Adamson & Regan, 1991; Bayley & Preston, 1996; Beebe, 1980; Weinberger, 1987; and Zampini, 1994). As sociolinguistic theory evolved, allowing for multiple factors in the discourse setting to influence speech (see Bell, 1984, for example), so too did research on variation in SLA. Features such as the ethnicity of the interviewer (Beebe & Zuengler, 1983), gender of the participant (Pica, Holliday, Lewis, Berducci, &

Newman, 1991), the topic of conversation (Selinker & Douglas, 1985), and the degree of cultural empathy shown by the interviewer (Berkowitz, 1989) were shown to correspond to variation in use. Other issues such as variation on a single task across time (Duff, 1993), the effects of task type and planning time (Bygate, 1999; Crookes, 1989; Robinson, 2001; Skehan & Foster, 1999), and the effects on negotiation of meaning (McDonough & Mackey, 2000) have also been studied. The range of factors that can influence learner language, include (at least): situational context, illocutionary meaning, linguistic contexts, discourse contexts, and planning conditions (Ellis, 1999). Nearly all of these relate to variation across tasks.

Moreover, a few studies have demonstrated differences across tasks in the types of structures that are elicited. In a reanalysis of data collected using a grammaticality judgment task, an oral interview, and an oral narration task from 20 learners of English, 10 with Arabic as a first language (L1) and 10 with Japanese an L1, Tarone and Parrish (1988) showed that learners produced different quantities of each of four types of noun phrases (NPs) on the oral tasks, even though some types of NPs were consistently more frequent than others regardless of the task. In addition, the accuracy for each type of NP varied within a single task and in some cases accuracy also varied across all three tasks. Salaberry and López-Ortega (1998) confirmed these findings and showed that the degree to which it is possible to avoid a structure interacted with learner responses. In this latter study the structures examined (the preterit/imperfect contrast, null subjects, and articles in L2 Spanish) were determined by context and both written tasks (a cloze and a multiple-choice task) were contextualized. In a related study, Geeslin (2006) analyzed data from 72 beginning and intermediate learners of Spanish and demonstrated that the types of contexts elicited in a picture-description task differed significantly from those elicited through a semi-structured interview such that the contextual features that predict use of *estar* appeared with differing frequencies across tasks. Thus, differing rates of use of a variant may be related to differences in the quantity of features that prompt that structure.

Given the relatively large number of studies on task variation in an L2, it is somewhat surprising that more is not known about how such variation influences the elicitation of variable structures in Spanish. Our current investigation is not an examination of the characteristics of the two tasks included in the present study nor an evaluation of why they are different from each other. Instead, we seek to learn more about the type of data elicited through the WCT and the interview task because these two tasks are the most prevalent in the relevant literature. In so doing, we are able to connect our findings to the previous research on copula and mood choice in Spanish.

1.2 Copula Contrast in L2 Spanish

As stated previously, research on sociolinguistically-variable structures, such as the copula contrast in Spanish, has moved away from findings based on an analysis of errors and toward an analysis of the frequency of selection of a given variant and the linguistic contextual predictors of that selection (see Geeslin, 2005, for details). Thus, early studies on the stages of acquisition of the full range of functions of *ser* 'to be' and *estar* 'to be' (Ryan & Lafford, 1992; VanPatten, 1987) have given way to more detailed analyses of specific functions, such as the [copula + adjective] structure. Consequently, we now know that development across proficiency levels can be described in terms of the frequency of use of a given copula and the relative importance of the predictors of that use. For example, the frequency of use of *estar* increases across time. Likewise, the contextual feature 'susceptibility to change', which distinguishes referent + adjective combinations that are changeable from those that are not, is a more important predictor of language use at early stages of development than at later stages (Geeslin, 2000). It is also known that advanced learners and NSs in the United States show statistically similar rates of selection of the two copulas but differ in the contextual features that predict copula selection (Geeslin, 2003). At present, studies that employ these methods of analysis exist for learners of Spanish with a variety of L1s (see Geeslin & Guijarro-Fuentes, 2006, for Portuguese and Geeslin & Guijarro-Fuentes, 2005, for French, German, and English), for a range of learning contexts (see Geeslin, 2000, for instructed Spanish, Geeslin, 2003, for instructed Spanish with study abroad, and Geeslin & Guijarro-Fuentes, 2006, for naturalistic learning), and for a wide range of NS comparison groups (see Cortés-Torres, 2003, for New Mexico; Díaz-Campos & Geeslin, 2004, for Venezuela; Geeslin & Guijarro-Fuentes, 2007, for Galicia; Geeslin & Guijarro-Fuentes, forthcoming, for Catalonia, Valencia, and the Basque Country; and Ortiz López, 2000, for Puerto Rico). In summarizing our knowledge to date, the frequency of use of *estar* is increasing over time, and certain

linguistic features, such as predicate type, the degree to which an attribute is susceptible to change, and the class of the adjective are relatively consistent predictors of *estar* use across groups (for more on these variables see section 3.4). It should be clear that this variability cannot be described in terms of contexts as defined by a group of lexical items because context must be understood as a combination of one category of each of the relevant syntactic, semantic, and pragmatic predictive variables.

In contrast to the range of research contexts in which copula choice has been examined, relatively few elicitation methods have been used in such investigations. Geeslin (2000) elicited data from beginning and intermediate learners of Spanish using a semi-structured interview, a picture-description task, and a WCT and found significant differences in the frequency of use of *estar* across tasks as well as differences in the predictors of that use from one task to another. Nevertheless, that is the only study that analyzes the predictors of copula choice in the [copula + adjective] structure that has used multiple means of data elicitation. Thus, research of this type on NSs has been limited to either a WCT or an interview (with the exception of Ortiz López, 2000, who used both but did not distinguish between the two in his analysis). To date, no analysis of the differences in the data elicited by these two tasks for NSs of Spanish exists. Likewise, research on advanced learners has been limited to the WCT, with no other means of triangulation available. The evidence for differences across tasks found for beginning and intermediate learners (Geeslin, 2000; 2006), suggests that research on other populations will also show similar effects. More importantly, a comprehensive analysis of the differences that exist across elicitation tasks is essential in understanding how the results from one study compare to another. Thus, we take as a point of departure the need to reconcile the gap between methods most commonly used to investigate copula choice for NSs and the methods that have been used to date to investigate advanced L2 learners.

1.3 Mood Choice in L2 Spanish

While a variety of tasks have been used to research L2 mood choice (e.g., see Collentine, 1995, Kornuc, 2004, and Lubbers Quesada, 1998, for interviews; Collentine, 1997, for a computerized generation task; Collentine, 1995, for a controlled-oral production activity; Stokes, Krashen, & Kartchner, 1998, for an oral sentence-completion task; Kornuc, 2004, for written sentence and dialogue-completion tasks; Gudmestad, 2006, for a WCT; and Gudmestad, 2005, for essays), the study of this structure has only recently shifted away from an analysis of errors to an analysis of the frequency of choice of a given variant (indicative or subjunctive) and the linguistic predictors of that choice. Thus, only Gudmestad (2006), in her comparison of intermediate and advanced-level learners' mood selection in traditional subjunctive contexts on a WCT has employed the methods of analysis necessary to incorporate sociolinguistic variation[2] into the research design. As with previous research on copula choice, we share the assumption with sociolinguists that predictors of use and rates of use are the key to analyzing this type of variation.

Even without limiting our review to those studies that incorporate sociolinguistic variation into the design, research comparing mood choice across tasks is limited. Kornuc (2004) compared accuracy rates of subjunctive and indicative use on three tasks across three proficiency levels. She equated task difficulty with level of control and ranked the tasks from the least difficult/most controlled to the most difficult/least controlled: a written sentence-completion task, a written dialogue-completion task, and an oral interview. She expected accuracy to decrease as task difficulty increased/level of control decreased, but her hypothesis was only partially confirmed. The traditional subjunctive items supported the hypothesis but the traditional indicative items did not. Kornuc suggested that the unexpected results regarding the traditional indicative verbs were due to the participants' higher proportional use of familiar verbs (e.g. *ser*) in the interview than in the written tasks. These findings suggest that the relationship between mood choice and task difficulty/level of control remains unclear. Kornuc's work also demonstrates that task variation has not been examined in conjunction with the more recent analytical method of identifying mood-choice predictors. Likewise, no comparable study exists for NS mood choice across tasks.

2. The Current Study

As shown in the preceding review, a key component of generalizing findings across studies, and

indeed of understanding the SLA of copula and mood choice in Spanish, is knowing how the interview and the WCT compare. Thus, our goal is to examine the differences across these two tasks for both NSs and advanced NNSs. In so doing, we learn more about the SLA of these two variable structures, and about the relationship that each elicitation task has to L2 research. The current study was guided by a single research question: How do NSs and advanced NNSs of Spanish each vary their choice of the copulas and the subjunctive and indicative moods from a WCT to an interview task? To answer this question completely, each of the following issues were addressed: (1) Does the frequency of choice vary across tasks? (2) Do the predictors of choice vary across tasks? (3) Does NS choice differ from NNS choice?

2.1 Participants

The participants in the current study were 20 NSs and NNSs of Spanish. The NSs (N=10) were from Argentina, Chile, Colombia, Ecuador, Mexico, Puerto Rico, Spain, and Uruguay. Seven were female and three were male, and they ranged in age from 24 to 37. The NNSs of Spanish (N=10) were native English speakers. They had studied Spanish in various countries: Argentina, Colombia, Costa Rica, Ecuador, Mexico, and Spain. They had spent an average of 13.9 months abroad and had studied Spanish for an average of 7.56 years. They were evenly divided between genders and ranged in age from 22 to 30. All participants were graduate students in Hispanic linguistics or literature and were instructors of undergraduate Spanish courses.

2.2 Elicitation Tasks

The participants completed a background questionnaire, a language proficiency test, a sociolinguistic interview, and two WCTs. The preceding section was based on information from the background questionnaire. The language proficiency test was a multiple-choice activity integrated into a story written in Spanish, consisting of 25 items covering various grammatical concepts. Each item had three options, and participants selected the appropriate response. Next, the participants completed a 30-minute sociolinguistic interview with two NSs of Spanish (one female and one male). The interview questions were the same for all speakers and included topics such as their plans for winter vacation, why the United States seems to be such a divided country, and if NSs or NNSs of a language make better language teachers. The interviews were digitally recorded and later transcribed.

Lastly, the participants completed two WCTs, one for copula choice and the other for mood choice. The participants received a packet with the two WCTs together; roughly half of them completed the copula-choice WCT first while the other half began with mood-choice WCT. A WCT is comprised of a series of paragraph-length contexts that build upon a single story. Following each context is an item that is integrated into the story as dialogue. Within each item there are two sentences, and the participants are asked to select between three options, deciding whether they prefer sentence A or sentence B or if they like both. The copula-choice WCT was comprised of 28 items for *ser*/*estar*. The two sentences following each context differed only in that one sentence used *ser* and the other used *estar*. Similarly, the two sentences following each context in the mood-choice WCT differed only in the mood of the verb in the dependent clause (indicative or subjunctive). The task consisted of 35 items. Each WCT was balanced for their respective independent linguistic variables (to be discussed shortly). The order of *ser* and *estar* and the order of the subjunctive and the indicative changed throughout each task. In other words, one variant was not always used in sentence A and the other always in sentence B. An example of an item from each WCT is available in Appendix A. As stated earlier, these tasks were selected for the present investigation because they are the elicitation tasks that exist in the literature, not because of specific characteristics of either.

2.3 Coding Scheme

Two separate analyses were conducted for each of the dependent variables examined in the current study: copula choice and mood choice. We begin with a description of the independent contextual predictors associated with each structure prior to describing the analysis itself. For copula choice, each [copula + adjective] structure included on the WCT or produced in the interview task in

non-past contexts[3] was identified as a single token. The dependent variable for each was the copula chosen. The independent variable which is the focus of the current study was task type, distinguishing those tokens elicited on the WCT from those elicited through the interview. Additionally, in order to answer the questions raised about the predictors of copula choice across groups and across tasks, we coded for several other independent variables. These variables have been shown in previous research to best describe the copula contrast and language change. The first such variable, predicate type, distinguishes between those contexts that are limited in time (stage-level) from those that are unlimited or ongoing (individual-level) (Carlson, 1977; Fernández Leborans, 1999; Leonetti, 1994). The variable susceptibility to change distinguishes referent + adjective combinations that are changeable from those that are not (Silva-Corvalán, 1986; 1994b). The variable frame of reference distinguishes those referents that are compared to a group of like objects from those that are compared to themselves at another moment in time (Delbecque, 1997; Falk, 1979). The final linguistic variable, adjective class, divides adjectives according to their semantic class and includes categories such as size, age, and physical appearance (Vañó-Cerdá, 1982). These variables, their respective categories, and examples of each are shown in Table One.

Table 1. Linguistic Variables for Copula Choice

Variable	Categories	Example
Predicate type	[- Individual]	*Hoy, Elena está enferma* 'Today, Elena is sick'
	[+ Individual]	*Elena es simpática* 'Elena is nice'
Susceptibility to change	[+ Susceptible]	*El niño es pequeño* 'The boy is small'
	[- Susceptible]	*El coche es pequeño* 'The car is small
Frame of reference	[+ Comparison]	*El niño está alto* 'The boy is (grew) tall'
	[- Comparison]	*El niño es listo* 'The boy is smart'
Adjective class	[Age]	*Joven* 'young'
	[Size]	*Grande* 'large'
	[Physical appearance]	*Gordo* 'fat'
	[Description / evaluation]	*Difícil* 'good'
	[Desc. of a person(ality)]	*Inteligente* 'intelligent'
	[Color]	*Azul* 'blue
	[Mental state]	*Animado* 'animated'
	[Physical state]	*Cansado* 'tired'
	[Sensory characteristic]	*Sabroso* 'tasty'
	[Status]	*Casado* 'married'

The linguistic variables described in Table One were used in conjunction with several social variables such as age, gender, country of origin (for the NSs), location of study abroad (for the NNSs), length of time abroad (for the NNSs), and language proficiency test score. These variables were the same for the analyses of copula and mood choice.

Turning to the dependent variable mood choice, all dependent clauses produced in the interview and included on the WCT in non-past contexts were analyzed as a single token. The dependent variable was the mood (subjunctive or indicative) chosen. Similar to copula choice, the independent variable of task type differentiates between the data elicited from the WCT and from the interview. The independent linguistic variables included to examine the predictors of mood choice across tasks and across groups were semantic category, verb morphology, and futurity. The variable semantic category analyzes mood choice according to various semantic meanings; such as volition, uncertainty, comment, and assertion; conveyed in matrix clauses (e.g., Blake, 1981; Collentine, 1995; Gudmestad 2005, 2006; Kornuc, 2004; Lubbers Quesada, 1998; Lynch, 2000; Silva-Corvalán, 1994a). The variable verb morphology distinguishes between irregular present tense, regular present tense, and past tense verbs in the dependent clause (Collentine, 1997; Gudmestad, 2006; Lubbers Quesada, 1998), and the variable futurity differentiates between future and non-future discourse contexts (Gudmestad, 2006). The variables, their categories, and examples are provided in Table Two.

Table 2. Linguistic Variables for Mood Choice

Variable	Categories	Example/Description
Semantic category (of the matrix clause)	[Assertion]	*Sé que va/vaya a nevar.* 'I know that it is going [INDIC/SUBJC] to rain.'
	[Comment]	*Es importante que uno se alimenta/alimente bien.* 'It's important that one eats [INDIC/SUBJC] well.'
	[Uncertainty]	*No creo que mi papá te puede/pueda ayudar.* 'I do not think that my dad can [INDIC/SUBJC] help you.'
	[Volition]	*Quiero que mis padres me visitan/visiten.* 'I want that my parents visit [INDIC/SUBJC] me.'
	[Other]	*Limpio la casa los domingos porque hay/haya mucho desorden.* 'I clean the house on Sundays because there is [INDIC/SUBJC] a lot of mess.'
Verb morphology	[Regular]	Verbs whose stems are the same in the present indicative and the present subjunctive (e.g., *estudia/estudie* [INDIC/SUBJC] 'studies')
	[Irregular]	Verbs whose stems differ between the present indicative and the present subjunctive (e.g., *tiene/tenga* [INDIC/SUBJC] 'has')
	[Past]	Past-tense verbs used in a non-past context (e.g., *estuvo/estuviera* [INDIC/SUBJC] 'was')
Futurity	[Future]	Dependent clauses produced in a future context (e.g., future plans, etc.)
	[Non-future]	Dependent clauses that are not produced in a future context (e.g., opinions of the town in which the participants attend school)

As previously mentioned, the following independent variables were also included in the analysis of mood choice: age, gender, country of origin (for the NSs), and location and length of study abroad (for the NNSs), and language proficiency test score.

2.4 Analysis

For the purpose of statistical analysis, the dependent variables in the WCTs were collapsed into binary variables (this task allowed for the selection of both variants). The selection of choosing both variants was collapsed with the selection of *estar* and with the selection of the indicative mood since sociolinguistic research has shown that the use of these variants is spreading (see also Geeslin, 2005). In order to answer the first subcomponent of our research question, cross tabulations were carried out for each structure to show the frequency of selection of the dependent variables across tasks and across participant groups. Chi-square (X^2) tests were also conducted to learn whether the distribution of the dependent variables was significantly different across tasks and across participant groups. Binary logistic regressions were carried out to answer the second sub-question. The regression analyses identified the independent variables that predicted NS and NNS choice of each dependent variable. A qualitative comparison of the significant predictors of copula and mood choice across tasks and across participant groups was subsequently conducted in order to answer the third sub-question.

3. Results

As stated previously, each participant completed a proficiency test, a sociolinguistic interview, and two WCTs. We begin this section with a brief summary of the results on the proficiency test. Subsequently, the results on the other two tasks will be grouped according the dependent variables (copula or mood choice) and according to the method of analysis (frequency of selection or predictors of selection). This section concludes with additional analyses to aid in the interpretation of the results that are less transparent.

Each participant completed a proficiency test and the results for both participant groups are summarized in Table Three. The table shows that while not all NSs received a perfect score, no NS scored below 88 percent. In contrast, only four of the ten NNS participants performed within the NS

range, and none of the NNSs earned a perfect score.

Table 3. Results for Native and Non-native Speakers on the Proficiency Test

Participant group	Native speakers	Non-native speakers
Range	88-100	64-92
Mean	95.2	81.6
Standard deviation	3.493	10.632

Among the questions that the current study was designed to answer was whether or not the frequency of choice of the copulas varied across tasks. The percentage of the time that *ser* and *estar* were chosen on each task is summarized in Table Four.

Table 4. Frequency of Copula Choice

Copula Choice	NS-WCT		NNS-WCT		NS-Interview		NNS-Interview	
	#	%	#	%	#	%	#	%
Estar	132	47.1	138	49.3	69	17.1	62	16.4
Ser	127	45.4	129	46.1	334	82.9	317	83.6
Both	15	5.4	13	4.6	0	0	0	0
Neither	6	2.1	0	0	0	0	0	0
Total	280	100	280	100	403	100	379	100

Note: NSs vs. NNSs on WCT (X^2=.115, df=1, Cramer's V=.014, p=.735), NSs vs. NNSs on interview (X^2=.081, df=1, Cramer's V=.010, p=.775), interview vs. WCT for NNSs (X^2=103.907, df=1, Cramer's V=.397, p=.000), and interview vs. WCT for NSs (X^2=.95.629, df=1, Cramer's V=.374, p=.000).

The data in Table Four and the corresponding X^2 tests demonstrate two important results. Firstly, the two participant groups were not significantly different from each other on either task. Whereas the NSs selected *estar* 47.1 percent of the time (132/280 tokens) on the WCT and 17.1 percent of the time (69/403 tokens) on the interview task, the NNSs selected *estar* with similar frequency (49.3% [138/280] and 16.4% [62/379], respectively). Secondly, both participant groups showed significant differences across tasks; they chose *estar* more frequently on the WCT than in the interview. In other words, NSs and advanced NNSs did not differ significantly from each other in the rates of choice of the copulas, but the interview and the WCT were significantly different from each other for both participant groups.

In order to determine whether the frequency of mood choice varied across tasks, a second set of cross tabulations and X^2 tests were carried out. Table Five shows the results. Both groups selected the subjunctive more frequently on the WCT than they used it in the interview. Similar to the results for copula choice, the X^2 tests indicated that the difference in frequency of mood choice across tasks was significant for both groups. The NSs chose the subjunctive 61.6 percent of the time (215/349 tokens) on the WCT and 18.4 percent of the time (311/1688 tokens) in the interview, while the NNSs chose the subjunctive 49.1 percent of the time (172/350 tokens) on the WCT and 6.6 percent of the time (75/1141) in the interview. However, unlike the results for *ser* and *estar*, X^2 tests revealed a significant difference between participant groups on each task. The NSs chose the subjunctive significantly more than the NNSs on both tasks.

Table 5. Frequency of Mood Choice

Mood Choice	NS-WCT		NNS-WCT		NS-Interview		NNS-Interview	
	#	%	#	%	#	%	#	%
Subjunctive	215	61.6	172	49.1	311	18.4	75	6.6
Indicative	125	35.8	164	46.9	1377	81.6	1066	93.4
Both	9	2.6	14	4.0	0	0	0	0
Total	349	100	350	100	1688	100	1141	100

Note. NSs vs. NNSs on WCT (X^2=10.981, df=1, Cramer's V =.125, p=.001), NSs vs. NNSs on interview (X^2=81.150, df=1, Cramer's V =.169, p=.000). Interview vs. WCT for NNSs (X^2=351.168, df=1, Cramer's V =.485, p=.000), Interview vs. WCT for NSs (X^2=81.150, df=1, Cramer's V =.372, p=.000).

Differences across tasks and across participant groups were also examined in terms of the predictors of the dependent variable. These predictors are those social and linguistic variables included in the predictive model produced by a regression analysis test. An X indicates the variables that are included in the regression model and asterisks indicate the level of statistical significance of each. An X without an asterisk signifies a variable that is included in the model but is not significant. The results of the four regression analyses run for copula choice are summarized in Table Six with details in Appendix B.

Table 6. Predictors of Copula Choice

Variable	NS-WCT	NNS-WCT	NS-Interview	NNS-Interview
Predicate Type			X	
Susceptibility to Change			***X	**X
Frame of Reference	*X	***X		
Adjective Class	***X	***X	X	***X
Age		***X		
Gender		*X		
Proficiency	*X			
Country			X	

Note. *p<0.05, **p<0.01, ***p<0.001.

Table Six indicates that on the WCT the linguistic predictors of copula selection were frame of reference and adjective class. Individual frame of reference and the adjective classes color, mental state, physical state, and sensory characteristic favored *estar* for both groups. Additionally, the predictive model for the NSs on the WCT included the social variable proficiency (the score on the proficiency exam), and for the NNSs the social variables age and gender were included in the predictive models for the WCT. All social variables will be addressed below.

For the interview, the predictive factors for each group were different from those for the WCT. Only the linguistic factor adjective class remained in the predictive models for both groups (mental and physical state favored *estar*). For the NSs the linguistic factors predicate type and susceptibility to change were added to the predictive model. Likewise, the linguistic feature susceptibility to change was added to the predictive model for the NNSs. When predicate type was included in the predictive model, stage-level predicates favored *estar*, and, when susceptibility to change was included in the model, contexts that were susceptible to change favored *estar*. In examining the social predictors of copula choice, the NS model also included country of origin in the predictive model for the interview tasks, whereas the NNS model did not include any social variables. As with the results for frequency of selection, the similarities across groups were stronger than the similarities for either group across tasks.

Four separate regression analyses were also conducted to examine the predictors of mood choice. Table Seven illustrates these results with details in Appendix B.

Table 7. Predictors of Mood Choice

Variable	NS-WCT	NNS-WCT	NS-Interview	NNS-Interview
Semantic category	***X	***X	***X	***X
Verb morphology			**X	X
Futurity	***X	***X	***X	
Age		***X	***X	
Gender				
Proficiency		*X	***X	*X
Country/abroad		***X	*X	*X

Note. *p<0.05, **p<0.01, ***p<0.001.

A comparison between groups of the predictors on the WCT show that the same linguistic factors predicted mood selection: semantic category and futurity. Subjunctive selection was highest in the semantic category of volition, followed by comment, uncertainty, other, and lastly assertion. Both groups showed variation in all semantic categories with one exception: the NSs selected the subjunctive categorically in contexts of volition. For both groups, subjunctive selection was higher in future contexts than in non-future contexts. No social variables were included in the predictive model for the NSs, but age, gender, proficiency, and location of study abroad were significant predictors for the NNSs.

The regression models for the interview reveal that semantic category and verb morphology were included in the predictive models for both groups but that futurity was only a significant predictor for the NSs. The frequency of subjunctive use descended across categories in the same order as for the WCT, and NSs showed variable use in all categories whereas NNSs showed categorical use of the indicative in contexts of assertion. For both groups, subjunctive use was higher with regular verbs than with irregular verbs or past verbs in non-past contexts and, for NSs, higher in future than in non-future contexts. Furthermore, age, proficiency, and country of origin predicted NSs' use, and proficiency and location of study abroad predicted NNSs' use.

We conclude this section with details for the social variables in order to interpret their inclusion in the predictive models. For copula choice, age and gender were included in the model for NNSs on the WCT and, for mood choice, age and proficiency were included in the model for NNSs on the WCT and proficiency was predictive for NNSs in the interview. Further analyses revealed that all of the males were in the older age group, that 80 percent of the females were in the younger age group, and that the older age group, composed largely of males, had studied abroad significantly longer and had significantly higher scores on the proficiency test. Thus, the differences between ages and genders were actually a reflection of proficiency level and time abroad. Age was also a significant predictor of mood choice for NSs in the interview. The older age groups used the subjunctive less frequently than the younger groups, but they had also been living in the United States longer than the younger groups and all but one of the older speakers was married to a native English speaker. These findings suggest that the older speakers' higher use of the indicative may be a result of extended contact with English.

The remaining social factors included in the regression models were likely due to the small participant pool. Firstly, the location of study abroad was a significant predictor of mood choice for the NNSs on the WCT and in the interview. However, the NNSs had lived in a total of six countries, and four of them had studied in at least two different countries. Secondly, country of origin of the NSs was included in the regression models for copula and mood choice in the interview, yet Spain is the only country that had more than one speaker. Finally, proficiency was a significant predictor of copula choice on the WCT and of mood choice in the interview for NSs. It seems unlikely that the inclusion of this variable in the regression model is indeed a reflection of different proficiency levels, since all NSs scored 88 percent or higher on the language proficiency test. Instead, it appears that the result for proficiency can also be attributed to the small participant pool.

4. Discussion

The current study was guided by a single research question with three subcomponents. Each of these subcomponents will be addressed in the discussion that follows and our findings will be

connected to previous research where possible. In assessing whether or not the frequency of choice of the structures under examination varies across tasks, we can answer quite definitively that rates of choice vary. This result was true for both copula and mood choice and for both NSs and NNSs. Specifically, *estar* and the subjunctive were selected significantly more frequently on the WCT by both groups. This result alone has important implications. Sociolinguistic research on both copula and mood choice has used rates of use of these structures to assess the degree to which language change is taking place in a particular speech community. Likewise, frequency of choice of *estar* and of the subjunctive can also serve as indicators of proficiency level among L2 learners. Our results show that comparisons must not be made across tasks or research studies that employ different means of data elicitation.

Secondly, we sought to examine whether the predictors of choice vary across tasks. Our results show that the predictors of *estar* and of the subjunctive mood are different for the WCT and for the interview task. There are at least two issues to consider in light of these results. Firstly, predictors of use of a given structure are a valuable tool for comparing learner and NS data across tasks and research should look to include both frequency and additional measures when examining such differences. The second implication of these findings is one that raises additional questions about the nature of the differences between these two elicitation tasks. Geeslin (2006) demonstrated that a key difference between the interview task and a picture-description task was that the two tasks elicited different types of tokens (see also Tarone & Parrish, 1988). In light of that finding, one question is whether the differences we have found across tasks are the result of differing patterns of choice resulting from features of the task itself or from the fact that the types of tokens included on each task are inherently different in terms of the categories of the linguistic variables that predict choice and their distribution across tasks. Thus, the differences we found in predictors of copula and mood choice provide evidence of the differences across tasks and also provide tools for examining these differences more closely in the future.

Finally, we sought to address whether or not NSs differed from NNSs on the two elicitation tasks included in our study. Unlike the previous two issues addressed in this discussion, we found different results for the two structures examined on this third sub-question. For copula choice, there were no significant differences between NSs and NNSs on either task. In contrast, the two groups did differ significantly on both tasks for mood choice. Because previous research has focused on either copula or mood choice, rather than on both within the same learner population, this result is of particular interest because it serves to connect two distinct bodies of research in the field of SLA. Our finding suggests that copula choice is further along in the process of SLA than mood choice for these advanced learners. Although this type of conclusion may have been made in earlier research, that research was based on an error analysis and did not reflect what is known about the variable nature of these two structures. Such variation means that having reached a minimum threshold of use of a structure (e.g., 90% accuracy in 'obligatory' contexts) did not actually indicate whether or not overall rates of use fell within NS ranges of use. A second important implication of this final finding is that both tasks are effective means of eliciting learner data. We can conclude this because both the interview and the WCT served to distinguish NSs from NNSs (in cases where this was appropriate) and this distinction was consistent between measures of frequency and the predictors of choice. Thus, despite the differences in the results across tasks, the findings based on both appear to be quite valid.

5. Conclusions and Future Directions

The current study was designed to shed light both on the SLA of variable structures in Spanish and on task variation. Regarding the former, we have added to the existing evidence that NSs vary in copula choice and also have shown that NNSs are performing within the NS range of use for that structure. In other words, when variation is taken into account, advanced NNSs choose copulas in ways that are similar to NSs. In contrast, we have shown that NNSs at the advanced level do not yet choose subjunctive or indicative mood in the same way that NSs do. Thus, by including two variable structures in a single design, we have added to the work conducted on each by showing that the copula contrast appears to be acquired before the mood contrast in L2 Spanish. In examining the importance of our findings for research on task variation, we have shown that not only does such variation exist across the WCT and interview, but such variation influences both frequency of choice and the

predictors of choice of a given variant. Moreover, the elicitation task is a greater source of variability than the contrast between NSs and advanced NNSs. This means that our previous view of what speakers do was to some degree incomplete because a range of elicitation methods had yet to be employed.

In light of these conclusions, it can also be argued that there is important research yet to be conducted. For example, a closer inspection of individual contexts across tasks would allow for a more direct comparison of use across the WCT and the interview task. This would require researchers to compare only those contexts where each of the categories for all linguistic predictors are identical and would shed light on whether or not the key differences lie in the actual selection of a variant or in the types of tokens elicited by these two tasks (i.e., how many of the contexts are comparable across tasks?). Another important line of research will involve an examination of tasks that fall between the two tasks examined in the current study in regard to the continuum of controlled vs. free elicitation. For mood choice in particular, future research should further explore the category 'other' of the variable semantic category in order to identify specific trends within this group of lexical items. Finally, an expanded participant pool would allow further exploration of the individual effects found in the current study, such as those linked to country of origin for the NS group and the overlapping variables of age, gender, proficiency, and time abroad for our NNSs. In sum, the current study has both shed light on existing research and identified a path for the continuation of such inquiry.

Notes

1. The words 'choice' and 'chose' are employed in our discussion of the results in order to accurately describe both tasks at the same time. While on the interview task, terms such as 'use' would also be possible, this may not be appropriate for the WCT where learners do not actually produce an utterance, but rather they 'select' between options that are supplied for them.
2. Sociolinguistic research has shown that speakers exhibit variable use across a range of semantic categories (see Section 4.3 for more detail on this variable) and that some varieties of U.S. Spanish are undergoing change toward an increased use of the indicative (see for example, Blake, 1981; Lynch, 2000; and Silva-Corvalán, 1994a).
3. In order to maintain a consistent analysis, only non-past contexts were examined in the interviews for both copula and mood choice because past contexts were excluded from the design of the WCTs.

Appendix A. Sample WCT items, Copula Choice (translated to English) and Mood Choice (originally presented in English)

Paula and Raúl are planning to go out to a restaurant tonight. 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: Would you like to go in my car?

     A. Raúl: ¡Ay! ¡Qué bonita estás!   [How pretty you are!]   ___ I prefer sentence A.
     B. Raúl: ¡Ay! ¡Qué bonita eres!   [How pretty you are!]   ___ I prefer sentence B.
                                               ___ I like both A and B.

Isabel and Xavier have met for lunch to discuss where they will go on vacation over the summer. Isabel is interested in traveling to a different country. She has brought a couple of travel books with her and shows them to Xavier. She says:
A. "Quiero que viajamos en Italia, en España o en México." [I want to travel in Italy, Spain or Mexico.]
B. "Quiero que viajemos en Italia, en España o en México." [I want to travel in Italy, Spain or Mexico.]
   ___ I prefer sentence A.   ___ I prefer sentence B.   ___ I like both.

Appendix B

Details of Regression Analyses

Regression Analyses	N	Model X^2	df	p	-2Log likelihood	Nagelkerke R^2	% predicted
Cop-NS-WCT	280	164.53	10	0	222.932	0.593	83.6
Cop-NNS-WCT	280	183.047	11	0	203.385	0.641	83.9
Cop-NS-INT	403	274.985	18	0	94.009	0.825	94.8
Cop-NNS-INT	379	244.302	10	0	93.442	0.806	96.8
Mood-NS-WCT	349	251.728	5	0	213.117	0.698	88.8
Mood-NNS-WCT	350	244.205	12	0	240.895	0.67	86.3
Mood-NS-INT	1688	405.132	16	0	1207.802	0.347	86.8
Mood-NNS-INT	1141	152.8	12	0	405.778	0.324	94.7

References

Adamson, Douglas, & Regan, Vera. (1991). The acquisition of community speech norms by Asian immigrants learning English as a L2. *Studies in Second Language Acquisition, 13*(1), 1-22.

Bayley, Robert, & Preston, Dennis. (1996). *SLA and linguistic variation*. Amsterdam: John Benjamins.

Beebe, Leslie. (1980). Sociolinguistic variation and style shifting in SLA. *Language Learning, 30*(2), 433-447.

Beebe, Leslie, & Zuengler, Jane. (1983). Accommodation theory: An explanation for style shifting in L2 dialects. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 195-213). Rowley, MA: Newbury.

Bell, Alan. (1984). Style as audience design. *Language in Society, 13*(2), 145-204.

Berkowitz, Diana. (1989). The effects of cultural empathy on L2 phonological production. In M. Eisenstein (Ed.), *The dynamic interlanguage: Empirical studies in L2 variation* (pp. 101-114). New York: Plenum Press.

Blake, Robert. (1981). Some empirically based observations on adult usage of the subjunctive mood in Mexico City. In J. Lantolf & G. B. Stone (Eds.), *Current research in Romance languages* (pp. 13-22).

Bygate, M. (1999). Quality of language and purpose of task: Patterns of learners' language in two oral communication tasks. *Language Teaching Research, 3*, 185-214.

Carlson, Gregory. (1977). A unified analysis of the English bare plural. *Linguistics and Philosophy, 1*, 413-456.

Chaudron, Craig. (2003). Data collection in SLA research. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 762-828). Malden, MA: Blackwell Publishing.

Cortés-Torres, Mayra. (2004). ¿Ser o estar? La variación lingüística y social de estar más adjetivo en el español de Cuernavaca, Mexico. *Hispania, 87*, 788-795.

Collentine, Joseph. (1995). The development of complex sentences and mood-selection abilities by intermediate-level learners of Spanish. *Hispania, 78*(1), 122-135.

Collentine, Joseph. (1997). The effects of irregular stems on the detection of verbs in the subjunctive. *Spanish Applied Linguistics, 1,* 3-23.

Crookes, Graham. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition, 11*, 367-383.

Delbecque, Nicole. (1997). The Spanish copulas ser and estar. In M. Verspoor, K. Dong Lee, & E. Sweetser (Eds.), *Lexical and syntactical constructions and the construction of meaning* (pp. 247-270). Amsterdam: John Benjamins.

Díaz-Campos, Manuel, & and Geeslin, Kimberly L. (2004). Copula use in the Spanish of Venezuela: Social and linguistic sources of variation. Paper presented at New Ways of Analyzing Variation, Ann Arbor, MI.

Dickerson, Leslie, & Dickerson, Wayne. (1977). Interlanguage phonology: Current research and future directions. In P. Corder & E. Roulet (Eds.), *Actes du 5ème Colloque de Linguistique Appliqué de Neuchâtel: The notions of simplification, interlanguages and pidgins and their relation to second language acquisition* (pp. 18-29). Geneva: Droz.

Duff, Patricia. (1993). Tasks and interlanguage performance: An SLA research perspective. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 57-95). Bristol, PA: Multilingual Matters LTD.

Ellis, Rod. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition, 9*, 1-20.

Ellis, Rod. (1999). Item versus system learning: Explaining free variation. *Applied Linguistics, 20*(4), 460-480.

Falk, Johan. (1979). *Ser y estar con atributos adjetivales*. Uppsala: Alqvist and Wiksell.

Fernández Leborans, María J. (1999). La predicación: Las oraciones copulativas. In I. Bosque & V. Demonte

(Eds.), *Gramática descriptiva de la lengua española* (pp. 2354-2460). Madrid: Espasa.

Geeslin, Kimberly 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 CLASP* (pp. 50-66). Somerville, MA: Cascadilla Press.

Geeslin, Kimberly L. (2003). A comparison of copula choice in advanced and NS Spanish. *Language Learning, 53.* 703-764.

Geeslin, Kimberly L. (2005). *Crossing disciplinary boundaries to improve the analysis of second language data: A study of copula choice with adjectives in Spanish.* Munich: LINCOM Europa Publishers.

Geeslin, Kimberly L. (2006). Linguistic contextual features and variation in L2 data elicitation. In C. Klee & T. Face (Eds.), *Selected proceedings from the 7th CLASP (*pp. 74-85). Somerville, MA: Cascadilla Press.

Geeslin, Kimberly L., & Guijarro-Fuentes, Pedro. (2005). The acquisition of copula choice in instructed Spanish: The role of individual characteristics. In D. Eddington (Ed.), *Studies in the acquisition of the Hispanic languages: Papers from the 6th CLASP* (pp. 66-77). Somerville, MA: Cascadilla Press.

Geeslin, Kimberly L., & Guijarro-Fuentes, Pedro. (2006). The second language acquisition of variable structures in Spanish by Portuguese speakers. *Language Learning,* 56(1), 53-107.

Geeslin, Kimberly L., & Guijarro-Fuentes, Pedro. (2007). Spanish in contact with Galician: Social factors, language use, language learning and the case of Copula choice. In K. Potowski & R. Cameron (Eds.), *Edited volume on Spanish in contact with other languages.* Amsterdam: John Benjamins Publishers.

Geeslin, Kimberly L., & Guijarro-Fuentes, Pedro. (forthcoming). Variation in contemporary Spanish: The role of bilingualism and language contact in the changing status of *estar. Bilingualism, Language and Cognition.*

Gudmestad, Aarnes. (2005). The L2 development of the Spanish subjunctive according to semantic categories. Paper presented at the Second Language Research Form, New York City, New York.

Gudmestad, Aarnes. (2006). L2 Variation and the Spanish subjunctive: Linguistic features predicting mood selection. In C. Klee & T. Face (Eds.), *Selected proceedings from the 7th CLASP (*pp. 170-184). Somerville, MA: Cascadilla Press.

Kornuc, Sandra Posada. (2004). L2 use and development of mood selection in Spanish complement clauses (Doctoral dissertation, University of Southern California, 2004). *Dissertation Abstracts International, 64,* 9.

Labov, William. (1994). *Principles of linguistic change: Internal factors.* Oxford: Blackwell.

Leonetti, Mario. (1994). Ser y estar: estado de la cuestión. *Barataria,* 1, 182 - 205.

Lubbers Quesada, Margaret. (1998). L2 acquisition of the Spanish subjunctive mood and prototype schema development. *Spanish Applied Linguistics*, *2*, 1-23.

Lynch, Andrew. (2000). The subjunctive in Miami Cuban Spanish: Bilingualism, contact, and language variability (Doctoral dissertation, University of Minnesota, 2000). *Dissertation Abstracts International, 60,* 8.

McDonough, Kim, & Mackey, Allison. (2000). Communicative tasks, conversational interaction and linguistic form: An empirical study of Thai. *Foreign Language Annals, 33*, 82-92.

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.

Pica, Teresa; Holliday, Lloyd; Lewis, Nora; Berducci, Dom; & Newman, Jean. (1991). Language learning through interaction: What role does gender play? *Studies in Second Language Acquisition, 13*, 343-376.

Robinson, Peter. (2001). Task complexity, task difficulty and task production: Explaining interaction in a componential framework. *Applied Linguistics,* 22, 27-57.

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.

Salaberry, Rafael, & López-Ortega, Nuria. (1998). Accurate L2 production across language tasks: Focus on form, focus on meaning and communicative control. *Modern Language Journal, 82*(4), 514-532.

Selinker, Larry, & Douglas, Dan. (1985). Wrestling with 'context' in interlanguage theory. *Applied Linguistics, 6*, 190-204.

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

Silva-Corvalán, Carmen. (1994a). The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change*, *6*(3), 255-272.

Silva-Corvalán, Carmen. (1994b). Language contact and change: Spanish in Los Angeles. Oxford: Claredon.

Skehan, Peter. & Foster, Pauline. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, *49*, 93-120.

Stokes, Jeffery; Krashen, Stephen; & Kartchner, John. (1998). Factors in the acquisition of the present subjunctive in Spanish: The role of reading and study. *ITL: Review of Applied Linguistics*, *121-122*, 19-25.

Tarone, Elaine. (1988). *Variation in Interlanguage*. London: Edward Arnold.

Tarone, Elaine, & Parrish, Betsy. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning*, *38*, 21-43.

Vañó-Cerdá, Anontio. (1982). *Ser y Estar + Adjetivos*. Tubingen: Narr.

VanPatten, Bill. (1987). The acquisition of ser and estar: Accounting for developmental patterns. In B. VanPatten, T. Dvork, & J. Lee (Eds.), *Foreign language learning: A research perspective* (pp. 61-75). New York:

Newbury House.

Weinberger, Steven. (1987). The influence of linguistic context on syllable simplification. In G. Ioup & S. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a L2 sound system* (pp. 401-417). Cambridge: Newbury House Publishers.

Zampini, Mary. (1994). The role of native language transfer and task formality in the acquisition of Spanish spirantization. *Hispania,* 77, 470-481.

Selected Proceedings of the
10th Hispanic Linguistics Symposium

edited by Joyce Bruhn de Garavito and Elena Valenzuela

Cascadilla Proceedings Project   Somerville, MA   2008

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 Aarnes Gudmestad. 2008. Comparing Interview and Written Elicitation Tasks in Native and Non-native Data: Do Speakers Do What We Think They Do? In *Selected Proceedings of the 10th Hispanic Linguistics Symposium*, ed. Joyce Bruhn de Garavito and Elena Valenzuela, 64-77. Somerville, MA: Cascadilla Proceedings Project.

or:

Geeslin, Kimberly L. and Aarnes Gudmestad. 2008. Comparing Interview and Written Elicitation Tasks in Native and Non-native Data: Do Speakers Do What We Think They Do? In *Selected Proceedings of the 10th Hispanic Linguistics Symposium*, ed. Joyce Bruhn de Garavito and Elena Valenzuela, 64-77. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1786.