L2 Variation and the Spanish Subjunctive: Linguistic Features Predicting Mood Selection

Aarnes Gudmestad
Indiana University

Native English speaking second language (L2) learners of Spanish typically begin learning indicative and subjunctive mood selection early in classroom instruction, yet the subjunctive is one of the many grammatical concepts that take a long time for learners to acquire (see for example, Terrell et al. 1987). Although the subjunctive has been investigated in second language acquisition (SLA) research, few studies have examined the morphological and semantic features that predict mood selection in L2 learners (Collentine 1995, Lubbers Quesada 1998). Consequently, the purpose of the present study is to examine subjunctive selection in a written preference task among intermediate level and advanced level L2 learners of Spanish. In order to better understand the development of the Spanish subjunctive mood among L2 learners, it aims to examine whether certain morphological and semantic linguistic features predict subjunctive selection and to compare the results between two proficiency levels.[1]

1. Background

Much of the research on the Spanish subjunctive and L2 learners has focused on reading and comprehension (Lee 1987, 1998, Lee & Rodriguez 1997, Stokes et al. 1998), simplified input (Collentine 1997a, Leow 1993, 1995), processing instruction (Collentine 1998, Farley 2001a, 2001b, 2004a, 2004b), the development of complex syntax (Collentine 1995, 1997b, Collentine et al. 2002), and accuracy (Stokes 1988, Stokes & Krashen 1990, Terrell et al. 1987). Lubbers Quesada's (1998) analysis focused on morphological and semantic features predicting L2 mood selection. Collentine (1995) also addressed this issue.

First, Collentine (1995) examined subjunctive use in noun clauses among intermediate level learners, all native speakers of English.[2] This group of learners (N=38) carried out a controlled oral production activity. The task was comprised of 44 drawings accompanied by glossed images and a caption. The participants listened to a question for each drawing and then gave a tape recorded oral response. An example of one task item included a picture of two people working together. The caption read: *María: 'Carlos, no vendiste casi nada el mes pasado.' Carlos: 'Entiendo, trabajaré más horas este mes.'* 'María: "Carlos, you sold almost nothing last month." Carlos: "I understand, I will work more hours this month."' (Collentine 1995:133). The participants were then asked questions such as: *¿Qué están haciendo los dos empleados?* 'What are the two employees doing?' or *¿Qué quiere María?* 'What does María want?' (Collentine 1995:127). While the former could be answered with a single clause, the latter would be answered correctly with a noun clause requiring the subjunctive. Fifty questions were asked: 30 elicited noun clauses (half requiring the subjunctive and half the indicative) and 20 were distracters. Results concerning semantic linguistic features revealed that

---

[1] The term 'linguistic feature' is used in the current study to refer to characteristics both at the sentence level and in the textual context of each item that are associated with the subjunctive.

[2] Two groups of learners completed two different tasks in Collentine (1995). The summary presented here is limited to the data from one task because it is the only one in which semantic linguistic features were analyzed. The reader is directed to Collentine (1995) for a complete discussion of the results for the two tasks.

learners produced subjunctive forms at different rates according to the semantic category. Learners used the subjunctive correctly in 41% of the contexts of influence (reports of commands and volition), followed by 36% of the contexts of doubt/denial and finally 23% of the contexts of emotion.

In a related investigation, Lubbers Quesada (1998) analyzed the variability of subjunctive use in interviews of intermediate level L2 learners of Spanish. These participants were enrolled in an intensive summer abroad program in Mexico in order to complete their university language requirement. An examination of subjunctive contexts at the sentence level revealed that three linguistic features predicted subjunctive use: irregular subjunctive verbs, expressions of desire and expressions of futurity. Specifically, matrix verbs expressing desire and those expressing futurity promoted the use of the subjunctive. The absence of some of these features resulted in a lower use of the subjunctive. The two feature combinations with the lowest subjunctive use were: (1) the absence of futurity and irregular subjunctive verbs with the presence of desire and (2) the absence of futurity and desire. Moreover, the presence of irregular subjunctive verbs promoted subjunctive use when the independent clause had the following feature combinations: (1) futurity without desire, (2) desire without futurity and (3) the absence of desire and futurity. Lubbers Quesada concluded that as L2 learners acquire the Spanish subjunctive, they establish a different system from that of adult native Spanish speakers. While syntactic, morphological and semantic features characterized the L2 system identified for these learners, adult native speakers use pragmatic information in mood selection (see for example, Lunn 1989).

Thus, Collentine (1995) and Lubbers Quesada (1998) suggested that sentence level linguistic features are relevant in subjunctive use among intermediate level L2 learners. Further inquiry in this area could add to their contributions by examining other linguistic features, exploring other means of data collection and analyzing L2 development among learners of different language proficiency levels.

2. The current study

The present investigation is largely motivated by Lubbers Quesada (1998) and is different from her work in five ways. First, subjunctive contexts in Lubbers Quesada were identified by matrix clauses that required the subjunctive according to the Spanish textbook the learners were using in their classes. Subjunctive contexts are not, however, solely determined at the sentence level; contextual information also affects mood selection. As a result, the current study expanded the analysis of sentence level linguistic features as predictors of the subjunctive to include contextual linguistic features.

Second, Lubbers Quesada (1998) classified vowel stem changing verbs, such as *volver* 'to return' as irregular, even though the indicative and the subjunctive stems are identical in the present tense. For example, *vuelve* 'she/he returns' is the indicative third person singular form and *vuelva* 'she/he returns' is the first and third person singular form. Following Collentine (1997a), vowel stem changing verbs were not classified as irregular in the present investigation because they undergo the same vowel change as regular verbs in the two moods; only the final, unstressed vowel differs. In the present study, verbs coded as irregular are limited to those whose present indicative and present subjunctive stems are different. Examples are *tener* 'to have' and *ser* 'to be.' The third person singular form of *tener* is *tiene* in the present indicative and *tenga* in the present subjunctive, and the third person singular form of *ser* is *es* in the present indicative and *sea* in the present subjunctive.

Third, in addition to continuing the investigation of the linguistic features of irregular subjunctive verbs, expressions of futurity and expressions of desire, the current study added expressions of emotion as a potential predictor of subjunctive selection. Fourth, data were collected with a written instrument instead of oral interviews. Finally, rather than studying one proficiency level, the present investigation examined intermediate and advanced level learners.

These differences have influenced the research questions and design of the current investigation on the present subjunctive.[3] The following research questions guide the present study:

---

[3] Because the current study only examined the present subjunctive, it will henceforth be referred to as the subjunctive.

(1) How frequently do English speaking adult intermediate level and advanced level learners of Spanish select the subjunctive in a written preference task?

(2) Do particular linguistic features predict intermediate and advanced level learners' selection of the subjunctive? When considered together, which factors predict subjunctive selection?

(3) If linguistic features predict the selection of the subjunctive, are the same features selected across proficiency levels?

3. Method

3.1 Participants

Two intact university L2 Spanish classes participated in the current study. All participants were native speakers of English. The intermediate level learners (N=17) were enrolled in the final course of the undergraduate language requirement (fourth semester) and had had instruction on the subjunctive. Thus, these learners appear to have been at a similar instructional stage as the participants in Lubbers Quesada (1998). Although no claims can be made about the subjunctive instruction participants had received in high school, in this university language program, subjunctive instruction begins as early as the second semester. In this fourth semester course, participants had already been instructed on the present subjunctive in evaluative statements, adverbial clauses, adjectival clauses and noun clauses of desire and influence and on the imperfect subjunctive in 'if' clauses. Nine were female and eight were male and their undergraduate majors included communications, business and various fields in the humanities and social sciences. Their ages ranged from 18 to 22 years, with an average age of 19.1 years. The advanced level participants (N=20) were studying introductory applied linguistics in a fourth year course. The subjunctive had not been addressed in this class prior to participation in the present investigation, but it is reasonable to assume that they had studied the subjunctive extensively in other courses. In particular, an advanced grammar course, which is required for the major or minor and must be taken before fourth year content courses, covers subjunctive use extensively. Fourteen participants were female and six were male. Eighteen learners were pursuing undergraduate degrees in education; business; and various fields in the humanities, social sciences and sciences; and two were Master's students in the Hispanic linguistics program. Their ages ranged from 20 to 35 years, with an average age of 22.3 years.

3.2 Data collection

Each participant completed three activities. The tasks were not timed. Learners spent between 30 and 40 minutes completing all tasks.

3.2.1 Background questionnaire

Participants first completed a background questionnaire to ensure common language learning experiences and personal information. They provided information on exposure they had had with Spanish and other languages, such as coursework, foreign travel and contact with native speakers. Personal information included age, gender, ethnicity, college major and parental education levels.

3.2.2 Language proficiency test

A language proficiency test, which served to confirm that the language abilities of the two groups were different from each other, followed the background questionnaire. This multiple choice cloze passage in Spanish consisted of 11 items covering a variety of grammatical concepts such as object pronouns, verb tense/aspect distinctions and verb conjugations. Each blank had three options and learners were asked to select the correct response. Descriptive statistics (Table 1) show that the intermediate level learners answered correctly an average of 6.24 items out of the 11 total items. The number of correct responses ranged from 1 to 9. The advanced level group earned a mean score of

10.15 out of the 11 possible points and their scores ranged from 7 to 11. In order to determine whether the language proficiency test scores of the two groups were statistically different, a chi square test was also performed. The chi square test (presented at the bottom of Table 1) showed significant results (χ^2 = 568.20, df = 7, Cramer's V = 0.876, p < 0.001). Therefore, the two participant groups were statistically different from each other; the advanced level learners scored significantly higher on the language proficiency test than the intermediate level learners.

Descriptive statistics	Proficiency level	
	Intermediate	Advanced
Mean	6.24	10.15
Range	8	4
Min. score	1	7
Max. score	9	11
Std. deviation	2.583	0.964

Table 1: Language proficiency test scores according to proficiency level (χ^2 = 568.20, df = 7, Cramer's V = 0.876, p < 0.001)

3.2.3 Written preference task

For the final activity, participants completed a written preference task, adapted from previous research on the Spanish copula (see for example, Geeslin 2003). The format for this task was a series of contextualized items written in English that built on a single story line about college students planning a summer vacation. These 35 items were presented in English to eliminate other linguistic and comprehensibility factors that may have affected performance on the instrument. Two Spanish sentences, differing only in the mood of the verb in the subordinate clause (one in the present subjunctive and one in the present indicative), followed each segment. These Spanish sentences were integrated into the story as dialogue. The order of the subjunctive and indicative clauses within each item varied throughout the task. Participants were asked to read each item and then to select the sentence or sentences they preferred. An example of an item from the instrument is available in Appendix A.

The written preference task consisted of 20 subjunctive contexts, or contexts in which the subjunctive is possible in Spanish. Although the subjunctive is required in these contexts in Standard Spanish, they are referred to as possible subjunctive contexts in the present study because previous research has shown that mood selection varies among native speakers of Spanish (see for example, Blake 1982, Ocampo 1990, Silva-Corvalán 1994, Studerus 1995a, 1995b, Torres 1989). These items were coded for the presence or absence of four linguistic features: irregular subjunctive verbs, expressions of futurity, expressions of desire and expressions of emotion (the coding procedures will be discussed shortly). The 15 remaining contexts followed the same format as the subjunctive items (English contexts with two sentences of Spanish dialogue differing only in mood of the subordinate clause verb) but required the indicative in Standard Spanish and served as distracters.[4]

In order to ensure that the task was interpreted appropriately, a group of six native speakers of various Spanish varieties completed the task before it was administered to the participants. These speakers were from Chile, Colombia, Mexico, Spain, Uruguay and Venezuela. All but one were pursuing graduate degrees in Hispanic literature or linguistics. The final speaker was a college educated professional. They all agreed that the subjunctive was possible in the 20 subjunctive contexts and that the indicative was possible in the 15 indicative contexts. In other words, there were no discrepancies among this group of native speakers; they selected the subjunctive in 100% of the subjunctive contexts and the indicative in 100% of the indicative contexts.

---

[4] The indicative items were not included in the analysis because it was not possible to adhere to the same coding schema employed for the subjunctive contexts. Although it was not feasible to follow the coding combinations for expressions of futurity, expressions of desire and expressions of emotion, the irregular verb feature was represented. In the subordinate clauses of the indicative items, eight verbs were regular and seven were irregular.

A written preference task was chosen for data collection for two reasons. First, the data from this instrument were easily comparable because all participants were exposed to the same linguistic situations. Although the study of natural speech is valuable, the language one learner uses may be quite different from another. In other words, it is possible that learners produce very few subjunctive contexts, that they do not produce a variety of subjunctive contexts and/or that they do not produce the same subjunctive contexts as other participants. This heterogeneity could make comparability of the data across participants more challenging. Second, as previously discussed, sentential and contextual features are both used to identify subjunctive contexts. This instrument design permitted the researcher to control for the presence and absence of linguistic features at both levels. Despite these advantages, two points should be noted. First, since previous research has shown that learners' use of linguistic structures varies across elicitation tasks (see for example, Tarone & Parrish 1988 and Geeslin this volume), results from the current study are not necessarily generalizable to other means of data collection. Second, a written preference task does not provide data on the development of subjunctive production.

3.3 Data collection

All learners completed the three tasks as part of a class activity, although their participation in the study was voluntary. Whereas all participants in the advanced level group completed the same packet of questionnaires, the intermediate level group received two different versions. In addition to selecting the preferred sentence(s) for the written preference task, the participants in the first subgroup were asked to write down what they were thinking as they completed each item. The second subgroup completed a written preference task identical to that of the advanced level group and one additional activity. After the task, they completed a cloze passage in Spanish in which they were asked to write verbs in the preterit and the imperfect. The purpose of this supplemental activity was to balance task completion time with that of the first intermediate level subgroup. The first subgroup did not complete the cloze passage activity. The comparison of these different research groups is part of another study, so this issue is left to future research. The two intermediate level groups were analyzed together in the current study.

3.4 Data coding

All tokens were coded for four linguistic variables. A token was defined as a context in which the subjunctive was possible in Standard Spanish (as mentioned earlier, there were 20 tokens in the instrument). The dependent variable was whether or not the subjunctive was selected in these contexts. In order to conduct a regression analysis (to be discussed shortly), this variable must be binary. The preference for the sentence with the subjunctive verb and the preference for both sentences were grouped together in the statistical analyses, although the preference for both sentences was not selected frequently.[5] These data were grouped with the preference for the subjunctive because the response may indicate that learners recognized that the subjunctive was possible in those contexts. The other category for the dependent variable consisted of the selection of a preference for the sentence with the indicative verb in the subordinate clause.

Each of the 20 tokens was coded for the presence or the absence of four independent variables: irregular subjunctive verbs, expressions of futurity, expressions of desire and expressions of emotion. Specifically, the coding for the linguistic features was determined by the Spanish sentences and the English context immediately preceding the dialogue. An explanation of the coding for the independent variables for one task item is provided in Appendix A.

As stated earlier, only verbs whose stems differed between the present indicative and present subjunctive were classified as irregular verbs. Following Collentine (1997a), verbs that undergo the same stem change from the infinitive to both the present indicative and present subjunctive were

---

[5] The preference for both sentences was selected twice by the intermediate level group (out of 340 tokens) and 16 times by the advanced level group (out of the 400 tokens).

categorized as regular verbs. The presence of irregular subjunctive verbs will be referred to as [+irregular] and the absence of this linguistic feature will be referred to as [-irregular].

While the coding for the irregular subjunctive verbs appeared only in the Spanish dialogue, the coding for expressions of futurity was overtly provided solely in the English context. Lexical items and verb forms indicating futurity were never provided in the Spanish dialogue in order to create consistency across all items in the task. Items were coded for the presence of futurity expressions ([+futurity]) when the periphrastic or morphological future verb tenses or temporal adverbs were provided in the English context. When these features were absent ([-futurity]), the temporal reference of the contexts was the present.

Unlike irregular subjunctive verbs and expressions of futurity, expressions of desire and those of emotion appeared both in the English context and in the Spanish dialogue. Items in which the English context and the Spanish dialogue included expressions of volition or hope were coded for the presence of desire ([+desire]). When these expressions were missing from the items, the tokens were coded for the absence of the linguistic variable ([-desire]). Similarly, English contexts and Spanish dialogue conveying emotion or feelings experienced by one of the characters in the story were coded for the presence of emotion ([+emotion]). When emotion was not expressed in the items, the tokens were coded for the absence of expressions of emotion ([-emotion]).

As a means to provide consistency in the task design, the presence of each of the four linguistic features appeared in isolation and in combinations with other features. There were a total of 10 feature combinations and each combination was represented twice in the task, yielding the 20 subjunctive contexts. Appendix B illustrates these permutations. Two combinations were not possible. First, subjunctive contexts with expressions of desire inherently imply futurity, so the pairing of the absence of expressions of futurity and the presence of expressions of desire was not possible. Contexts with the presence of both expressions of desire and those of emotion were also not an option. Although the two linguistic features could appear in the same English context, they cannot co-appear in the same matrix clause at the sentence level.

3.5 Data analysis

The 20 subjunctive contexts were isolated for statistical analysis. All tokens were coded for the linguistic and social variables and entered into the Statistical Package for the Social Sciences 11.0 (SPSS). The first statistical test performed distributional cross tabulations and chi square analyses. These tests revealed which individual independent variables were significantly related to subjunctive selection when considered in isolation.[6] Following other researchers investigating variation in SLA (see for example, Adamson & Regan 1991, Berdan 1996, Díaz-Campos 2004, Geeslin 2003), a multivariate analysis was carried out. A binary forward stepwise regression analysis was performed for each participant group. This logistic regression analysis examined all independent variables in relation to each other and then selected the variables that best predicted the dependent variable. Variables that did not improve the prediction of subjunctive selection were not included in the predictive model produced by this test.

---

[6] A chi square test was chosen instead of a parametric test because the written preference task measures the frequency of subjunctive preference and not accuracy; thus, there is no maximum score. In addition to the need for a maximum score, parametric tests, such as an ANOVA, require interval or ratio measurement scales, whereas nonparametric chi square tests permit nominal measurement scales. The binary dependent variable in the present study is nominal (preference for the subjunctive or the indicative). Finally, a chi square test does not require tokens to be normally distributed. Due to the design of the written preference task, the independent variables of desire, futurity and emotion do not have an equal number of contexts in which each individual feature is present or absent (for example, of the 20 subjunctive contexts in the instrument, four are [+desire] and 16 are [-desire]).

4. Results
4.1 Frequency of subjunctive selection

It can be recalled that the language proficiency test confirmed that the two participant groups were at different proficiency levels; the advanced level learners performed significantly higher on the language proficiency test than the intermediate level participants. Thus, the results for each group are presented separately. Moreover, the results in Table 2 reveal that the advanced level learners selected the subjunctive more than the intermediate level learners. The advanced level learners selected the subjunctive in 77.8% of the possible subjunctive contexts (311/400 tokens) whereas the intermediate level learners selected the subjunctive in 59.4% of the possible subjunctive contexts (202/340 tokens).[7] The chi square test shows that this difference between the two groups is significant; the advanced level participants selected the subjunctive significantly more than the intermediate level learners ($\chi^2 = 29.036$, df = 1, Cramer's V = 0.198, $p < 0.001$).

| Proficiency level | Subjunctive selection | | | |
| | Yes | | No | |
	Tokens	%	Tokens	%
Intermediate	202	59.4	138	40.6
Advanced	311	77.8	89	22.3

$\chi^2 = 29.036$, df = 1, Cramer's V = 0.198, $p < 0.001$

Table 2: Overall subjunctive selection according to proficiency level (Note: N = 340 for the intermediate level group and N = 400 for the advanced level group)

Just as the results for overall subjunctive selection indicated a significant difference between proficiency levels, the chi square tests for the linguistic features also reveal differences between the two groups. The data for each linguistic feature were submitted individually to a chi square test and a Cramer's V test of effect size to yield the results for both the intermediate and advanced level learners. Tables 3 and 4 provide a summary of the chi square tests performed for each individual linguistic feature. Each row represents a separate statistical test and shows the chi square value, the degrees of freedom and the value of Cramer's V. The degree to which each test was significant is indicated with asterisks. The results from these tests indicate that the linguistic features significantly related to subjunctive selection differed between the intermediate and the advanced level learners. The only statistically significant linguistic feature related to subjunctive selection for the intermediate level group was irregular subjunctive verbs ($\chi^2 = 7.025$, df = 1, Cramer's V = 0.144, $p < 0.01$). Expressions of futurity, expressions of desire and expressions of emotion were not significant (Table 3).

Linguistic feature	χ^2	df	Cramer's V
Irregular verbs	7.025**	1	0.144**
Futurity	0.165	1	0.022
Desire	3.321	1	0.099
Emotion	0.165	1	0.022

Table 3: Chi square tests for linguistic features for intermediate level learners, **$p < 0.01$.

In contrast to the intermediate level group, all linguistic features were significant for the advanced level learners (Table 4). Irregular subjunctive verbs ($\chi^2 = 7.645$, df = 1, Cramer's V = 0.138, $p < 0.01$), expressions of futurity ($\chi^2 = 10.812$, df = 1, Cramer's V = 0.164, $p < 0.01$), expressions of desire ($\chi^2 = 19.784$, df = 1, Cramer's V = 0.222, $p < 0.001$) and expressions of emotion ($\chi^2 = 6.513$, df = 1, Cramer's V = 0.128, $p < 0.05$) were significantly related to subjunctive selection.

---

[7] It can be recalled that the token numbers for the two groups are different because there were more participants in the advanced level group (N=20) than in the intermediate level group (N=17).

Linguistic feature	χ^2	df	Cramer's V
Irregular verbs	7.645**	1	0.138**
Futurity	10.812**	1	0.164**
Desire	19.784***	1	0.222***
Emotion	6.513*	1	0.128*

Table 4: Chi square tests for linguistic features for advanced level learners, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$

The next four tables present the cross tabulations, or frequency of subjunctive selection, for each linguistic feature shown to be significant in the chi square tests; the observed and expected values are included. In Table 5 the data show that both proficiency levels selected the subjunctive more frequently when irregular subjunctive verbs were present. The intermediate level participants chose the subjunctive in 66.5% of the [+irregular] contexts (113/170 tokens). A comparison of the observed values with the expected values (59.4% or 101/170 tokens) also shows that subjunctive selection was higher than expected in [+irregular] contexts. Additionally, intermediate level participants selected the subjunctive in 52.4% of the [-irregular] contexts (89/170 tokens), and the observed value was lower than the expected value (59.4% or 101/170 tokens). Similarly, the advanced level learners selected the subjunctive in 83.5% of the [+irregular] contexts (167/200 tokens), yielding a higher observed than expected value (77.8% or 155.5/200 tokens). They selected the subjunctive less frequently in [-irregular] contexts; the subjunctive was chosen in 72.0% of these contexts (144/200 tokens). This value was also lower than the expected value (77.8% or 155.5/200 tokens). Thus, the observed and expected values from the cross tabulation, along with the chi square tests, indicate that the presence of irregular verbs was significantly related to subjunctive selection for intermediate and advanced level learners.

Irregular verbs	Proficiency level					
	Intermediate			Advanced		
	Tokens	Total tokens	%	Tokens	Total tokens	%
[+Irregular]						
Observed	113	170	66.5	167	200	83.5
(Expected)	(101)		(59.4)	(155.5)		(77.8)
[-Irregular]						
Observed	89	170	52.4	144	200	72
(Expected)	(101)		(59.4)	(155.5)		(77.8)

Table 5: Distribution of subjunctive selection according to irregular subjunctive verbs and proficiency level

Next, frequency data are presented for the linguistic features of futurity, desire and emotion for the advanced level learners. These results are not given for the intermediate level learners since the chi square tests indicated that these linguistic features were not significantly related to their subjunctive selection. Beginning with expressions of futurity, the frequency of subjunctive selection was higher when expressions of futurity were present. Table 6 reveals that advanced level learners selected the subjunctive in 83.3% of the [+futurity] contexts (200/240 tokens) and that the observed rate of subjunctive selection was higher than the expected value (77.8% or 186.6/240 tokens). Moreover, subjunctive selection was noticeably lower in [-futurity] contexts. The advanced level learners chose the subjunctive in 69.4% of these contexts (111/160 tokens). This observed value was lower than the expected value (77.8% or 124.4/160 tokens). Therefore, the chi square tests and observed and expected values from the cross tabulation demonstrate that the presence of expressions of futurity was significantly related to subjunctive selection for advanced level learners.

Futurity	Subjunctive selection		
	Tokens	Total tokens	%
[+Futurity]			
Observed	200	240	83.3
(Expected)	(186.6)		(77.8)
[-Futurity]			
Observed	111	160	69.4
(Expected)	(124.4)		(77.8)

Table 6: Distribution of subjunctive selection according to expressions of futurity for advanced level learners

Similar to irregular subjunctive verbs and expressions of futurity, advanced level learners chose the subjunctive more frequently when expressions of desire were present. Table 7 reveals that subjunctive selection was very high in [+desire] contexts for advanced level learners. They chose the subjunctive in 96.3% of these tokens (77/80 tokens), more frequently than the expected value (77.8% or 62.2/80 tokens). In [-desire] contexts they selected the subjunctive less frequently than expected; the observed value was 73.1% (234/320 tokens) and the expected value was 77.8% (248.8/320 tokens). Once again, the observed and expected values from the cross tabulation, along with the chi square tests, indicate that the presence of expressions of futurity was significantly related to subjunctive selection for the advanced level learners.

Desire	Subjunctive selection		
	Tokens	Total tokens	%
[+Desire]			
Observed	77	80	96.3
(Expected)	(62.2)		(77.8)
[-Desire]			
Observed	234	320	73.1
(Expected)	(248.8)		(77.8)

Table 7: Distribution of subjunctive selection according to expressions of desire for advanced level learners

In contrast to the other three linguistic features, subjunctive selection was lower when expressions of emotion were present. Table 8 shows that advanced level learners selected the subjunctive less frequently than expected in [+emotion] contexts. The expected value was 77.8% (124.4/160 tokens) and the observed value was 71.3% (114/160 tokens). Furthermore, subjunctive selection was higher when expressions of emotion were absent. The participants chose the subjunctive in 82.1% of [-emotion] contexts (197/240 tokens). This observed value was higher than the expected value (77.8% or 186.6/240 tokens). Consequently, when interpreted in light of the cross tabulation, the significant chi square result for expressions of emotion reveals that the absence of expressions of emotion was significantly related to subjunctive selection for the advanced level learners. This result may be attributed to the task design and will be discussed later.

Emotion	Subjunctive selection		
	Tokens	Total tokens	%
[+Emotion]			
Observed	114	160	71.3
(Expected)	(124.4)		(77.8)
[-Emotion]			
Observed	197	240	82.1
(Expected)	(186.6)		(77.8)

Table 8: Distribution of subjunctive selection according to expressions of emotion for advanced level learners

4.2 Predictors of subjunctive selection

In order to answer the research questions concerning the predictors of subjunctive selection, a regression analysis was performed for each proficiency level. As previously mentioned in section 3.5 on data analysis, a binary logistic regression identifies the independent variables that best predict the dependent variable. Table 9 illustrates the results for the two regression analyses; an X signifies the variables that are included in the regression model and asterisks indicate the statistical significance of the variables. The details of the statistical tests for the regression are presented in Table 10; the model χ^2, the degrees of freedom, the level of significance, the -2 log likelihood, the Nagelkerke R^2, and the percent of all tokens the model predicts are provided.

These results show differences between the two proficiency levels. First, only irregular subjunctive verbs predicted subjunctive selection for the intermediate level learners (Table 9). In other words, when the remaining linguistic features were added to the model after irregular subjunctive verbs were added, none of these variables contributed to the model to better predict subjunctive selection. Table 10 illustrates that this logistic regression model was statistically significant and explained 59.4% of the subjunctive selection variance for intermediate level participants.

However, for advanced level learners irregular subjunctive verbs and expressions of desire worked together to predict subjunctive selection (Table 9). Thus, when irregular subjunctive verbs and expressions of desire were included in the model, they better predicted subjunctive selection together than either feature alone. Adding expressions of futurity or expressions of emotion did not improve the predictive power of the model. This regression model for the advanced level learners explained more variation, or had more predictive power, than the model for intermediate level participants. Table 10 indicates that the model including irregular subjunctive verbs and expressions of desire was statistically significant and explained 77.8% of the subjunctive selection variance.

Linguistic feature	Intermediate	Advanced
Irregular verbs	X*	X*
Futurity		
Desire		X**
Emotion		

Table 9: Significant predictors of subjunctive selection according to proficiency level. (Note: X = variable included in the model, *$p < 0.01$. **$p < 0.001$)

Statistical tests	Proficiency level	
	Intermediate	Advanced
χ^2	7.053	34.084
df	1	2
p	< 0.01	< 0.001
-2 Log likelihood	452.167	389.959
Nagelkerke R^2	0.28	0.125
Percent predicted	59.4	77.8

Table 10: Details of statistical tests for the logistic regression models according to proficiency level

5. Discussion

The results will now be discussed in relation to the research questions. First, both participant groups selected the subjunctive in the written preference task, but the advanced level learners selected the subjunctive significantly more than the intermediate level learners. The data revealing that intermediate level participants selected the subjunctive in over half of the possible subjunctive contexts suggest that these learners have the linguistic ability to make mood selections. Furthermore, the results showing that the advanced level participants selected the subjunctive in over three fourths of the possible subjunctive contexts suggest that language development is taking place. As L2 learners progress from an intermediate language course to a fourth year content course, their ability to select the subjunctive in possible subjunctive contexts improves.

Second, the data showed that the linguistic features predicting mood selection differed between the two groups. Irregular subjunctive verbs were the only significant factor for intermediate level participants. Although this result supports the findings of Lubbers Quesada (1998), it is interesting to note that the more stringent criterion for irregular verb classification following Collentine (1997a) did not change the significance of this linguistic feature. Moreover, unlike Lubbers Quesada (1998), in which linguistic features worked together to predict mood use for intermediate level learners, in the current study irregular subjunctive verbs did not work together with other features to predict subjunctive selection for the intermediate level group. The data also revealed that the presence of irregular subjunctive verbs, expressions of futurity and expressions of desire as well as the absence of expressions of emotion were significantly related to subjunctive selection for the advanced level group. However, only irregular verbs and expressions of desire worked together to better predict subjunctive selection in the regression model.

Since desire was a significant predictor of subjunctive use for the intermediate level learners in Lubbers Quesada (1998) and because learners in Collentine (1995) demonstrated higher accuracy rates of subjunctive production in contexts of influence (similar to desire in Lubbers Quesada 1998 and the current study) than in contexts of doubt or emotion, the results concerning expressions of desire in the present investigation are particularly interesting. A closer analysis of the [+desire] contexts reveals that subjunctive selection varied across task items for the intermediate, but not the advanced, level learners. The four [+desire] contexts included the following matrix clauses in Spanish: *quiero que* 'I want that,' *prefiero que* 'I prefer that,' *ojalá* 'I hope' and *espero que* 'I hope/wish that.' While intermediate level participants chose the subjunctive in 94.1% (16/17 tokens) of the contexts with *espero que*, subjunctive selection was noticeably lower with *quiero que* (58.8% or 10/17 tokens), *prefiero que* (58.8% or 10/17 tokens) and *ojalá* (64.7% or 11/17 tokens). In contrast, frequency of subjunctive selection was high in all contexts for the advanced level learners. They chose the subjunctive in all contexts of *quiero que* (100.0% or 20/20 tokens) and *espero que* (100.0% or 20/20 tokens), in 95.0% (19/20 tokens) of the *ojalá* contexts, and in 90.0% (18/20 tokens) of the *prefiero que* contexts. Although the current study supports the conclusion that morphological and semantic linguistic features are relevant to the L2 development of the subjunctive, this discussion on expressions of desire seems to suggest that analyses of individual matrix clauses are also warranted in tracing L2 subjunctive development.

Moreover, task design may explain the results for expressions of futurity for the advanced level participants. It can be recalled that overt expressions of futurity were only provided in the context, not in the dialogue sentences. Since this linguistic feature was a significant predictor for advanced level learners but was not included in the regression model, it appears that these learners have begun to use this contextual information for subjunctive selection. However, it also seems that this specific linguistic feature has not yet reached the level of significance of the two features included in the regression model that were also provided at the sentence level (irregular subjunctive verbs and expressions of desire).

In addition, the advanced level group data for the expressions of emotion were particularly interesting because the absence of this linguistic feature significantly predicted subjunctive selection. Task design may also explain this result; some tokens were coded as [+desire] and [-emotion]. Expressions of desire were included in the regression model predicting subjunctive selection but expressions of emotion were not. Thus, it is possible that the significant result for [-emotion] is actually an indicator of the strength of the [+desire] feature; expressions of desire seem to be a stronger predictor of mood selection than expressions of emotion.

Sociolinguistic research on the Spanish subjunctive offers another possible explanation for the results concerning expressions of emotion. For example, Silva-Corvalán (1994) investigated subjunctive use among three generations of Mexican-American bilinguals living in Los Angeles. An analysis of interview data revealed that mood use was highly variable in matrix clauses of comment (similar to expressions of emotion in the present study). Since both the indicative and the subjunctive were used in these contexts, it seems possible that the input learners receive is also variable. If research shows that variation exists in the input learners receive, it may be that the subjunctive develops later among L2 learners in expressions of emotion than in other contexts in which native speakers exhibit less variation in subjunctive use.

Finally, the only linguistic feature that significantly predicted subjunctive selection for both proficiency levels was irregular subjunctive verbs. Semantic features of futurity, desire and emotion were only significant for advanced level participants. Even though the results for intermediate level learners in the current study differed from those in Lubbers Quesada (1998), the results for both proficiency levels supported her general conclusion that learners use morphological, syntactic and semantic features to identify subjunctive contexts. More specifically, the present investigation has suggested that morphological features of subjunctive verbs seem to predict subjunctive selection before other contextual features such as desire, futurity and emotion.

6. Conclusion

The current study showed that the linguistic features predicting mood selection in a written preference task differed between intermediate and advanced level L2 learners of Spanish. Only the presence of irregular subjunctive verbs was significantly related to subjunctive selection for intermediate level learners whereas the presence of irregular subjunctive verbs, expressions of futurity and expressions of desire and the absence of expressions of emotion was significantly related to subjunctive selection for advanced level learners. Furthermore, no linguistic features worked together with irregular subjunctive verbs to predict intermediate level subjunctive selection, but irregular subjunctive verbs and expressions of desire worked together to predict selection for advanced level learners.

Future research can take a number of directions. Clearly, the present investigation does not provide a complete understanding of mood selection; additional linguistic features such as expressions of doubt and expressions of evaluation need to be examined. Other studies employing a written preference task for data collection could examine the differences between sentence and context level linguistic features predicting subjunctive selection. An examination of the possible effect of the inclusion of other verb tenses (future, for example) into the subordinate clause of the Spanish test sentences may also be interesting, since the current study required a forced choice between two present tense moods. The results for the expressions of emotion suggest that the effects of variation in input on acquisition of the subjunctive could be investigated as well. Finally, in addition to investigating

additional proficiency levels, analyses of learner production data are necessary for a more complete understanding of interlanguage development. Production data gathered from interviews or essays could be analyzed for the presence or absence of linguistic features associated with the subjunctive and then submitted to statistical analyses to learn what linguistic features predict subjunctive use and if the results are comparable across elicitation tasks.

Appendix A: Sample task item and coding explanation

The following is a sample item from the written preference task. The English translations for the Spanish sentences are provided here for the reader but were not available to the participants.

1. 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 prefer sentence A.
   [I want that we travel in Italy, Spain or Mexico.]
B. "Quiero que viajemos en Italia, en España o en México."   ___ I prefer sentence B.
   [I want that we travel in Italy, Spain or Mexico.]

                                               ___ I like both.

The coding for this example is as follows: [-irregular] because *viajar* 'to travel' is a regular verb, [+futurity] because the discussion in the contexts centers on a vacation that has not yet happened ("they will go on vacation over the summer"), [+desire] because Isabel is stating her preference for a vacation destination in the context ("Isabel is interested in traveling") and in the matrix clause *quiero que* 'I want that' and [-emotion] because no feelings are expressed in the context or in the dialogue.

Appendix B: Linguistic feature combinations in subjunctive contexts
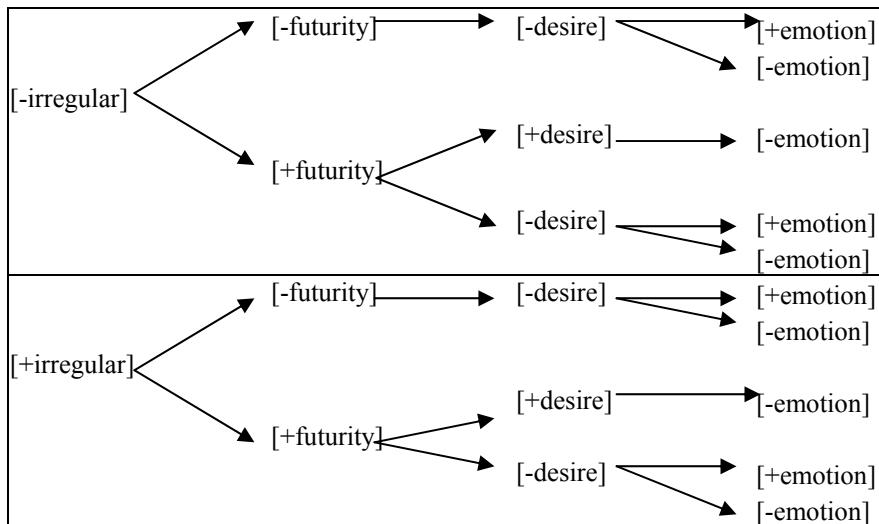
References

Adamson, Douglas, and Vera Regan. 1991. The acquisition of community norms by Asian immigrants learning English as a second language: a preliminary study. Studies in Second Language Acquisition 13.1-22.

Berdan, Robert. 1996. Disentangling language acquisition from language variation. Second language acquisition and linguistic variation, ed. by Robert Bayley and Dennis R. Preston, 204-44. Amsterdam: John Benjamins.

Blake, Robert. 1982. Some empirically based observations on adult usage of the subjunctive mood in Mexico City. Current research in Romance languages, ed. by James Lantolf and Gregory B. Stone, 13-22. Bloomington: Indiana University Linguistics Club.

Collentine, Joseph G. 1995. The development of complex syntax and mood-selection abilities by intermediate level learners of Spanish. Hispania 78.122-35.

Collentine, Joseph G. 1997a. The effects of irregular stems on the detection of verbs in the subjunctive. Spanish Applied Linguistics 1.3-23.

Collentine, Joseph G. 1997b. Pragmatics and the generation of syntactically complex utterances by foreign language learners of Spanish. Contemporary perspectives on the acquisition of Spanish, vol. 2: production, processing, and comprehension, ed. by William R. Glass and Ana Teresa Pérez-Leroux, 25-40. Somerville, MA: Cascadilla Press.

Collentine, Joseph G. 1998. Processing instruction and the subjunctive. Hispania 81.576-87.

Collentine, Joseph G.; Karina Collentine; Victoria Clark; and Eric Friginal. 2002. Subjunctive instruction enhanced with syntactic instruction. Structure, meaning, and acquisition in Spanish: papers from the 4th Hispanic Linguistics Symposium, ed. by James F. Lee, Kimberly L. Geeslin, and J. Clancy Clements, 32-45. Somerville, MA: Cascadilla Press.

Díaz-Campos, Manuel. 2004. Context of learning in the acquisition of Spanish second language phonology. Studies in Second Language Acquisition 26.249-73.

Farley, Andrew P. 2001a. Authentic processing instruction and the Spanish subjunctive. Hispania 84.289-99.

Farley, Andrew P. 2001b. Processing instruction and meaning-based output instruction: a comparative study. Spanish Applied Linguistics 5.57-94.

Farley, Andrew P. 2004a. Processing instruction and the Spanish subjunctive: is explicit information needed? Processing instruction: theory, research, and commentary, ed. by Bill VanPatten, 227-39. London: Lawrence Erlbaum.

Farley, Andrew P. 2004b. The relative effects of processing instruction and meaning-based output instruction. Processing instruction: theory, research, and commentary, ed. by Bill VanPatten, 143-68. London: Lawrence Erlbaum.

Geeslin, Kimberly L. 2003. A comparison of copula choice: native Spanish speakers and advanced learners. Language Learning 53.703-64.

Geeslin, Kimberly L. This volume. Task design, discourse context and variation in second language data elicitation.

Lee, James F. 1987. Comprehending the Spanish subjunctive: an information processing perspective. Modern Language Journal 71.50-7.

Lee, James F. 1998. The relationship of verb morphology to second language reading comprehension and input processing. Modern Language Journal, 82.33-48.

Lee, James F., and Renato B. Rodriguez. 1997. The effects of lexemic and morphosyntactic modifications on L2 reading comprehension and input processing. Contemporary perspectives on the acquisition of Spanish, vol. 2: production, processing, and comprehension, ed. by William R. Glass and Ana Teresa Pérez-Leroux, 135-57. Somerville, MA: Cascadilla Press.

Leow, Ronald P. 1993. To simplify or not to simplify: a look at intake. Studies in Second Language Acquisition 15.333-56.

Leow, Ronald P. 1995. Modality and intake in second language acquisition. Studies in Second Language Acquisition 17.79-90.

Lubbers Quesada, Margaret. 1998. L2 acquisition of the Spanish subjunctive mood and prototype schema development. Spanish Applied Linguistics 2.1-23.

Lunn, Patricia V. 1989. Spanish mood and the prototype of assertability. Linguistics 27.687-702.

Ocampo, Francisco. 1990. El subjuntivo en tres generaciones de hablantes bilingües. Spanish in the United States: sociolinguistic issues, ed. by John J. Bergen, 39-48. Washington, D.C.: Georgetown University Press.

Silva-Corvalán, Carmen. 1994. The gradual loss of mood distinctions in Los Angeles Spanish. Language Variation and Change 6.255-72.

Stokes, Jeffrey. 1988. Some factors in the acquisition of the present subjunctive in Spanish. Hispania 71.705-10.

Stokes, Jeffrey, and Stephen Krashen. 1990. Some factors in the acquisition of the present subjunctive in Spanish: a re-analysis. Hispania 73.805-6.

Stokes, Jeffrey; Stephen Krashen; and John Kartchner. 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.

Studerus, Lenard. 1995a. Mood variability in border Spanish. Bilingual Review 20.2.99-113.

Studerus, Lenard. 1995b. Some unresolved issues in Spanish mood use. Hispania 78.94-104.

Tarone, Elaine, and Betsy Parrish. 1988. Task-related variation in interlanguage: the case of articles. Language Learning 38.21-43.

Terrell, Tracy D.; Bernard Baycroft; and Charles Perrone. 1987. The subjunctive in Spanish interlanguage: accuracy and comprehensibility. Foreign language learning: a research perspective, ed. by Bill VanPatten, Trisha R. Dvorak, and James F. Lee, 19-32. Rowley, MA: Newbury House.

Torres, Lourdes. 1989. Mood selection among New York Puerto Ricans. International Journal of the Sociology of Language 79.67-77.

Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages

edited by Carol A. Klee and Timothy L. Face

Cascadilla Proceedings Project   Somerville, MA   2006

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:

Gudmestad, Aarnes. 2006. L2 Variation and the Spanish Subjunctive: Linguistic Features Predicting Mood
Selection. In *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First
and Second Languages*, ed. Carol A. Klee and Timothy L. Face, 170-184. Somerville, MA: Cascadilla
Proceedings Project.

or:

Gudmestad, Aarnes. 2006. L2 Variation and the Spanish Subjunctive: Linguistic Features Predicting Mood
Selection. In *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First
and Second Languages*, ed. Carol A. Klee and Timothy L. Face, 170-184. Somerville, MA: Cascadilla
Proceedings Project. www.lingref.com, document #1284.