Toward an Understanding of the Relationship between Mood Use and Form Regularity: Evidence of Variation across Tasks, Lexical Items, and Participant Groups

Aarnes Gudmestad
Virginia Polytechnic Institute and State University

Sociolinguistic research has shown that, for native speakers (NSs) of Spanish, mood distinction (the subjunctive-indicative contrast) is variable and undergoing change toward an increased use of the indicative (e.g., Silva-Corvalán, 1994) and that this variable use is conditioned by a range of factors, such as form regularity, semantic category, time reference, hypotheticality, and task (e.g., Gudmestad, 2010). Additionally, second-language (L2) research on Spanish has demonstrated that mood distinction is a late-acquired linguistic structure (e.g., Stokes, 1988) and that the acquisition of verbal moods can be characterized by the same variables that predict NS use (Gudmestad, forthcoming). Among the linguistic issues investigated in this body of sociolinguistic and L2 research is whether the subjunctive mood is linked more strongly with regular or irregular verbs (e.g., Collentine, 1997). Much, but not all, of this research has indicated that form regularity of the verb (in general, whether the verb in the mood-choice context is regular or irregular) is connected to NS and L2 mood distinction. The previous studies that have shown an effect for form regularity have resulted in diverse findings regarding the specific details of the relationship between the two variables. Thus, it remains unclear what role, if any, form regularity plays in mood distinction, yet this issue is an important one to address.

The quantitative paradigm has provided variationism with a powerful tool for understanding and predicting language variation and change, in part, because it allows researchers to analyze simultaneously linguistic and extra-linguistic factors conditioning language use (e.g., Bayley, 2002; Berdan, 1996). However, the effectiveness of a multivariate analysis (and arguably any kind of analysis) and the strength of the conclusions we draw from it hinge on an accurate and clear understanding of the variables under investigation. With such mixed results for the relationship between mood distinction and form regularity in Spanish, it seems necessary to examine the factor form regularity more thoroughly in order to account for the differences found across prior studies, to better operationalize this variable, and, ultimately, to advance our knowledge of mood distinction. In this way, the present investigation aims to clarify form regularity's role in the use of verbal moods in Spanish for learners and NSs by carrying out a detailed analysis of this linguistic factor.

1. Background

Research on the subjunctive-indicative contrast in Spanish has dealt with a range of topics including first-language acquisition (e.g., Pérez-Leroux, 1998), conversational implicature (e.g., Scott, 2004), descriptive accounts (e.g., Bosque, 1990), heritage speakers (e.g., Montrul, 2007), and processing instruction (e.g., Farley, 2004). However, the following literature review focuses only on the investigations that have analyzed the relationship between mood distinction and form regularity. Previous work on NSs of Spanish first showed that results for the connection between form regularity and mood use were inconclusive for Costa Rican speakers on written and oral correction tasks (Ross Veidmark & Umaña Aguiar, 1991). Later sociolinguistic research found that regular verbs predicted the subjunctive on an interview but not on a written contextualized task (WCT) for NSs in the United States (Geeslin & Gudmestad, 2008) and that, for a different group of NSs in the United States

(Gudmestad, 2010), form regularity predicted use, such that subjunctive use was highest with irregular verbs on three oral-elicitation tasks. Concerning L2 investigations, Collentine (1997) suggested that intermediate-level learners paid more attention to irregular than regular subjunctive verbs on a computer sentence-generation task, which increased the opportunity for irregular subjunctive forms to be incorporated into the interlanguage. Irregular subjunctive verbs predicted subjunctive use for intermediate-level learners on an interview task (Lubbers Quesada, 1998). Gudmestad (2006) demonstrated that irregular verbs influenced subjunctive selection on a WCT for intermediate- and advanced-level learners. In Gudmestad (forthcoming), form regularity predicted mood use on three oral-elicitation tasks for the two most advanced-level L2 groups only, with subjunctive use highest with irregular verbs. While these four L2 studies have identified a relationship between verb irregularities and the subjunctive, Geeslin and Gudmestad (2008) found conflicting results. Advanced learners used the subjunctive more frequently with regular verbs on an interview task (though this result was not found for the WCT).

What can be seen from this brief review is that there does not appear to be a uniform relationship between mood distinction and form regularity, but these inconsistent findings are not the only characteristic that distinguishes these studies from one another. Their research designs also differ, including notable differences among the participant populations, the data-elicitation instruments, and the operationalization of form regularity. For example, regarding the classification of form-regularity categories, most studies identified one regular and one irregular category. The category form-specific irregular in the current study corresponds to irregular in Collentine (1997), Gudmestad (2006) and Geeslin and Gudmestad (2008). In these three previous studies, what are regular and irregular verbs in the current study were grouped together as regular verbs. In Lubbers Quesada (1998), it appears that the categories form-specific irregular and irregular in the present investigation were classified into the same irregular category. Another difference is seen with Ross Veidmark and Umaña Aguiar (1991), who analyzed five categories of form-regularity. Similar to Collentine (1997), Gudmestad (2006), and Geeslin and Gudmestad (2008), the regular- and irregular-verb classifications in the current study were grouped together in their investigation of Costa Rican Spanish. In general, their remaining four categories distinguished among verbs included in the present investigation's form-specific irregular group. One difference among studies that impacts the classification of form regularity is that each of the prior investigations limited their analysis to the present tense, whereas Gudmestad (2010, forthcoming) and the present study included all tense/mood/aspect forms. In response to the incongruity of form-regularity classifications, Gudmestad (2010, forthcoming) and the current study distinguished between qualities of form regularity to a greater extent than studies with a dichotomous variable. However, the five categories set forth by Ross Veidmark and Umaña Aguiar (1991) could not be applied to the present investigation because the analysis was not restricted to the present tense (for definitions of the current study's three categories see the *Data Coding and Analysis* section).

In light of the diverse research methods and inconclusive findings across investigations, the purpose of the current study is to begin to reconcile these inconsistencies. It continues the analysis of the corpus and the tripartite categories of form regularity reported on in Gudmestad (2010, forthcoming) in order to begin to maintain some continuity across this area of research.[1] It initiates the process of better understanding the relationship between mood distinction and form regularity by exploring the role that tasks, lexical items, and participant groups may have in the varying results found across studies.

As previously noted, an examination of the body of work on form regularity and mood distinction quickly reveals that data have been collected from diverse tasks, including oral and written instruments and those that require producing language and others that do not. Because abundant evidence for task variation exists in the literature (e.g., Geeslin, 2010, Schilling-Estes, 2002, Tarone & Parrish, 1988) and because Gudmestad's (2010, forthcoming) learners and NSs varied their use of verbal moods across three oral-elicitation tasks, it seems reasonable to hypothesize that some of the varying effects found for form regularity may be due to the different instruments employed across studies. The current dataset also lends itself well to investigating task variation because all participants completed the same three tasks. Thus, the relationship between form regularity and mood use can be examined for each task individually and then the patterns of use can be compared across tasks.

---

[1] The current study recognizes that a range of factors influence mood distinction, even though its focus is on one of these variables.

Observations from sociolinguistic and L2 research appear to suggest that explorations of the lexical inventory of a dataset may also shed light on the relationship between form regularity and mood distinction. In her sociolinguistic investigation of speakers of Canadian French, Poplack (2001) found varying rates of subjunctive use across individual verbs and that suppletive/frequent verbs favored subjunctive use, while regular/rare verbs disfavored subjunctive use in French. Although there are differences in contexts of use and the morphological richness of verb endings between French and Spanish, it seems possible that Spanish exhibits varying patterns of subjunctive use across individual verbs too. Additionally, it has been suggested that learners' increasing vocabulary may impact the development of the copula in Spanish (Zyzik & Gass, 2008), so it may be the case that a developing lexicon influences the acquisition of other morphosyntactic structures as well. If these hypotheses are borne out in the data, then the differences found for the relationship between form regularity and mood use may be, at least in part, a reflection of the specific lexical items produced in each dataset (cf. Geeslin, 2006). Since the lexical items and their frequency of use are most likely different across datasets, the lexical inventory of datasets may be another source of variation across investigations. It is noted that this part of the present investigation is exploratory, since the data were not originally designed to analyze this issue.

Lastly, the participants studied in the existing literature include NSs with diverse backgrounds and learners with varying levels of proficiency in Spanish. The linguistic systems of these different participant groups are arguably not identical (cf. Ellis, 2008), so the populations investigated may be yet another source of variation. Therefore, if effects are found for the variables task and lexical item, it seems reasonable to investigate the similarities and differences of these effects across participant groups. The present corpus is well suited to examine this issue since the data come from five proficiency levels of learners and one group of NSs.

2. The Current Study

In order to more thoroughly understand the role that form regularity plays in variable mood use in Spanish, the current study was guided by the following research questions:
1. Do the effects of form regularity on mood use vary across data-elicitation tasks?
2. Within categories of form regularity, is there variability in mood use across lexical items (i.e., individual verbs)?
3. Do the effects found for task and lexical item correspond to learners, native speakers (NSs), or both?

2.1. Participants

Learners corresponding to five proficiency levels (N=130) and 20 NSs participated in the current study. Learners in Levels 1 through 4 were undergraduates and Level 5 was comprised of graduate-level speakers. Two intact classes of a third-semester, Spanish-language course constituted Level 1 (N=26); all participants were NSs of English. Level 2 (N=35) consisted of a fifth-semester, Hispanic-cultures course (two intact classes). One participant was a NS of Russian who learned English as a child and the others were NSs of English.

Levels 3 and 4 were comprised of learners in content courses beyond the fifth-semester of the program (two intact, third-year courses and two intact, fourth-year courses). Although Level 3 and Level 4 largely correspond to third-year and fourth-year courses, respectively, some adjustments based on previous coursework experience were made because students in this academic program have some flexibility with the sequence of courses after the sixth-semester of study. Learners in the third-year courses had completed an average of 3.75 courses (range=0-9 courses), whereas participants in the fourth-year courses had completed an average of 5.71 courses (range=2-9 courses). These means include the number of courses taken after the fifth-semester course, so all Level 3 and 4 participants had taken more Spanish courses than Level 2. To account for learners whose course sequence differed from the norm, learners in third-year courses who had taken six or more courses were regrouped to Level 4 and participants in fourth-year courses who had taken three courses or fewer were moved to Level 3. In Level 3 (N=26), one participant was a NS of English and German; the rest were NSs of English. Level 4 (N=23) consisted of two participants who were NSs of two languages (English and Dutch, English and Mandarin); the others were NSs of English.

Nineteen Level 5 participants were instructors of undergraduate Spanish courses; one was a graduate assistant (N=20). Two learners had already completed a graduate degree and 18 were graduate students. All were NSs of English.

The NSs of Spanish were graduate students, 17 of whom had experience teaching undergraduate Spanish courses. Their places of origin (Argentina, Chile, Colombia, Costa Rica, Dominican Republic, Guatemala, Mexico, Peru, Puerto Rico and Spain) reflect the diverse exposure to the Spanish language that the learners have likely received, either abroad or during their educational experiences in the United States. Additional information on the participants is available in Appendix A.

The language proficiency test (see *Data Collection* section) confirmed the use of course level to distinguish among proficiency levels. An ANOVA and a Post Hoc Scheffé examined differences in mean scores on the proficiency test between groups (Table 1) and revealed significant differences between groups with two exceptions. Significant differences in mean scores were not found between Levels 3 and 4 or among Level 4, Level 5, and the NSs. However, the average score for Level 3 was significantly different from the mean scores for Level 5 and the NSs, which indirectly identified a difference between Levels 3 and 4. The failure to distinguish between levels among the more proficient speakers is likely due to the ceiling effect for the test; the averages for Levels 4 and 5 and the NSs are close to the top score.

Table 1

Language Proficiency Test Scores

Group	*M*	*SD*	Min. score	Max. score
Level 1	4.96	1.64	2	8
Level 2	7.51	2.17	2	11
Level 3	8.96	2.31	4	11
Level 4	10.00	1.00	7	11
Level 5	10.75	0.55	9	11
NSs	10.80	0.41	10	11

Note. The highest possible score is 11. $F_{(5, 144)}=45.822$, $p<0.001$. A Post Hoc Scheffé identified four distinct groups.

2.2. Data Collection

The mood-use data came from three digitally-recorded, oral-production tasks in Spanish; an example of each is available in Appendix B.[2] In order to gain more information about the participants, they also completed a background questionnaire and a language proficiency test.

The first oral-elicitation task was a monologic role play (Task 1) and was an opportunity for extended language production in Spanish. Each of the six scenarios was followed by a series of questions, which were designed to elicit a variety of mood-choice contexts (see the *Data Coding and Analysis* section for a definition of mood-choice context). With every scenario the participants imagined themselves in a role and answered questions while pretending to speak to an imaginary person. The scenarios were presented on PowerPoint slides using a computer, and each one advanced automatically after the allotted time[3] (three minutes, 45 seconds).

---

[2] The data obtained from the oral-elicitation tasks were transcribed before they were coded and analyzed.

[3] The tasks were initially untimed, but the participants in the pilot study completed them in a wide range of time. In order to keep the task-completion time uniform, the tasks were timed for the data collection in the current study. The average time it took the pilot-study participants to complete each task was calculated and then this average time was evenly divided among items. Extra seconds were added to Tasks 2 and 3 so that the lower-level learners would have enough time to complete them. These two tasks were then piloted with third-semester learners to ensure that the allotted time was appropriate.

The format for the contextualized-clause-elicitation task (Task 2) and the contextualized-verb-elicitation task (Task 3) was similar. Both tasks consisted of a series of contextualized items (N=30 segments per task) that built on a single story, and the items represented a variety of mood-choice contexts. For Task 2 every segment was followed by the beginning part of a Spanish sentence that was integrated into the story as dialogue. Participants read each segment silently and the beginning part of the Spanish sentence aloud and then completed the sentence with a phrase that made sense in the context of the story. Many participants continued speaking after completing the sentence; these data were included in the analysis. For Task 3 each item was followed by a Spanish sentence that was integrated into the story as dialogue. In the mood-choice context of the sentence, an infinitival verb form was given in parentheses. Participants read each segment to themselves and then read the sentence aloud while providing the verb form they felt was appropriate in the context. For both tasks, the items were presented on a computer using PowerPoint and each slide advanced automatically (40 seconds per item on Task 2, 26 seconds per item on Task 3).[4]

The background questionnaire asked for personal information such as age and gender and details on their language experiences, including contact and foreign travel. The fourth instrument was a language proficiency tests that has shown to be a reliable measure for distinguishing proficiency levels (Woolsey, 2009) and was used in the present investigation to confirm the division of participant groups by course level. The 11-item, multiple-choice activity covered a range of grammatical structures in Spanish. Participants had three minutes to complete the test.

2.3. Data Coding and Analysis

The current study defines a token as a mood-choice context. 'Mood-choice context' is restricted to the following environments: verbs produced in subordinate clauses following an independent clause conveying a meaning of one of five semantic categories and verbs following a lexical expression that expresses one of five semantic categories. The five semantic categories are volition (e.g., *quiero que*, 'I want that'), comment (e.g., *es triste que*, 'it is sad that'), uncertainty (e.g., *tal vez*, 'perhaps'), temporality (e.g., *cuando*, 'when'), and assertion (e.g., *creemos que*, 'we believe that').[5] Detailed information regarding semantic category and its relationship to mood use is available in Gudmestad (2010, forthcoming).

The dependent variable is the verbal mood used in a mood-choice context: subjunctive or non-subjunctive. The term 'non-subjunctive' is used instead of 'indicative' because participants used non-finite forms in mood-choice contexts.[6] The two independent linguistic variables are form regularity and lexical item. Form regularity examines the verb produced in the mood-choice context. The three categories are regular, irregular, and form-specific irregular verbs. Regular verbs are those in which the stems of finite verb forms are identical to the stem of the infinitive or to the infinitive itself (e.g., *hablar*, 'to speak', infinitive; *habla,* present indicative; *hable,* present subjunctive; *habló,* preterit indicative; *hablara*, imperfect subjunctive). Irregular verbs undergo a vowel change in the stem of some (but not necessarily all) forms (e.g., *dormir*, 'to sleep', infinitive; *duerme,* present indicative; *duerma,* present subjunctive; *durmió,* preterit indicative; *durmiera*, imperfect subjunctive). Form-

---

[4] These oral-elicitation tasks were originally designed to enable form regularity, semantic category, time reference, and hypotheticality (independent linguistic variables) to be investigated systematically. The three tasks were also similar in that each provides contextual information beyond the sentence-level (i.e., participants do not produce or respond to a single sentence in isolation). Contextual information was needed in order to analyze the relationship between mood use and time reference and hypotheticality (two discourse-level variables). However, the format of each task was different because Task 1 offered the advantage of providing more spontaneous data, while Tasks 2 and 3 ensured that a large sample size could be obtained. See Gudmestad (2010, forthcoming) for more information on these instruments.

[5] These five categories were selected because previous research (e.g., Silva-Corvalán, 1994) has shown that they represent a range of variable use among NSs. Indicative use is high in contexts of assertion, both moods are used frequently in contexts of temporality and uncertainty, and subjunctive use is high in contexts of volition and comment.

[6] Verbs did not need to be target-like to be included in the analysis. Verb forms diverged from the target in terms of characteristics such as stress, stem, and verb-class ending and were identified as the verb form (i.e., conditional, preterit, etc.) they most closely resembled. These decisions were made after consulting with two researchers. When a verb form could not be classified, it was eliminated from the dataset.

specific irregular verbs have multiple verb forms (but not necessarily all) whose stem diverges from the infinitival stem. This deviation typically involves the addition of a consonant that is absent in the infinitive (e.g., *tener*, 'to have', infinitive; *tiene,* present indicative; *tenga,* present subjunctive; *tuvo,* preterit indicative; *tuviera*, imperfect subjunctive). The variable lexical item analyzes differences between individual verbs (e.g., *ser* 'to be' vs. *hacer* 'hacer' vs. *poner* 'to put', which are all form-specific irregular verbs). The two extra-linguistic independent variables are task and participant group. Task explores differences among the three data-elicitation tasks, participant group enables each learner and NS group to be investigated separately.

The data analysis began with cross-tabulations and chi-square (X^2) tests to examine the relationship between mood use and form regularity within each task, separately for each of the six participant groups. The chi-square tests revealed whether there was a significant relationship between the distribution of mood use and the independent variable form regularity when each task was considered in isolation. If a chi-square test was not significant, these results suggested that the distribution of verbal moods across the categories of form regularity was random.

The next step of the analysis was an exploration of the lexical inventory for a subset of the participant pool only. Because the examination of the relationship between form regularity and the variable lexical item was preliminary, the analysis focused on Levels 4 and 5 and the NSs, the three participant groups for whom the variable form regularity was a predictor of mood use (Gudmestad 2010, forthcoming). There were three parts to this exploratory analysis. Firstly, for each of the three participant groups, a type count of the lexical items (i.e., individual verbs) used in mood-choice contexts was conducted and their distribution across the categories of form regularity was assessed. Secondly, the frequency with which each lexical item was used and how these frequencies of use pattern across the form-regularity categories were examined. Since the current dataset was not originally designed to investigate vocabulary and since previous analyses had not been conducted on this issue for these data, the first two phases served to provide an initial characterization of the lexical inventory of the data. This description was considered a necessary step before moving on to an analysis that could lead to an answer to the second research question. Thirdly, in order to examine the relationship between mood use and lexical item, cross-tabulations and chi-square tests were carried out on a subset of the data, which included separate analyses for Levels 4 and 5 and the NSs. The chi-square tests showed whether or not there was a significant relationship between the distribution of mood use and the individual lexical items.

Finally, qualitative comparisons were made across participant groups. Similarities and differences were identified among the five proficiency levels of learners and the NSs with regard to the variable task and between Levels 4 and 5 and the NSs for the variable lexical item.

3. Results

Before examining the findings from the present investigation, results for the variables form regularity and task from Gudmestad (2010, forthcoming) that are pertinent to the current data analysis are reviewed briefly. The variable form regularity predicted mood use for Levels 4, 5 and the NSs, but not for Levels 1 through 3. The two most advanced proficiency levels and the NSs used the subjunctive most often with irregular verbs. Levels 4 and 5 produced the subjunctive least often with regular verbs, with the use of subjunctive slightly higher with form-specific irregular verbs. In contrast, NSs produced the subjunctive more often with regular than form-specific irregular verbs. In addition, the variable task predicted mood use for each of the six participant groups. All groups patterned similarly: subjunctive use was highest on Task 3 and lowest on Task 1.

3.1. Task

The presentation of the results begins with the results for the chi-square tests that examined the relationship between mood use and form regularity for each task and each participant group individually. Although form regularity predicted mood use for Levels 4 and 5 and the NSs only, the analysis includes all participant groups. Because previous research that employed different tasks found an effect for form regularity and because the variable task predicted mood use for each participant group, it may be possible that an effect for form regularity emerges on individual tasks before it

impacts mood use as a whole. Table 2 illustrates these results; an 'X' indicates a significant relationship between mood use and form regularity. No significant relationship was found between mood use and form regularity on Task 2 for any participant group. For Levels 1 and 5, a significant relationship between mood use and form regularity was not found on any of the three tasks. However, the distribution of mood use across categories of form regularity was significant on Task 1 for Level 2, Task 3 for Levels 3 and 4, and Tasks 1 and 3 for the NSs.

Table 2

Chi-square Tests for Mood Use and Form Regularity

Task	Participant group					
	Level 1	Level 2	Level 3	Level 4	Level 5	NSs
1		X				X
2						
3			X	X		X

Note. An 'X' indicates a significant result for the chi-square test.

The next series of tables provides the details of the cross-tabulations for the significant chi-square tests that examined the relationship between mood use and form regularity. Table 3 demonstrates that on Task 1 Level 2 used the subjunctive most often with form-specific irregular (5.8% or 18/313 contexts), followed by irregular (3.6% or 2/56 contexts), and, finally, regular (1.2% or 2/171 contexts) verbs (X^2=5.979, *df*=2, Cramer's V=0.105, p=0.050, 1 expected small cell).

Table 3

Distribution of Level 2's Subjunctive Use on Task 1 according to Form Regularity

Form regularity	#	%	Total
Form-specific	18	5.8	313
Irregular	2	3.6	56
Regular	2	1.2	171

Note. X^2=5.979, *df*=2, Cramer's V=0.105, p=0.050, 1 expected small cell.

Levels 3 (X^2=8.903, *df*=2, Cramer's V=0.109, p=0.012) and 4 (X^2=20.292, *df*=2, Cramer's V=0.173, p<0.001) showed a similar pattern of mood use on Task 3 (Table 4). Both groups produced the subjunctive most frequently with form-specific irregular (Level 3: 37.3% or 114/306 contexts and Level 4: 39.6% or 107/270 contexts), then irregular (Level 3: 32.8% or 83/253 contexts and Level 4: 34.6% or 80/231 contexts.), and, lastly, regular (Level 3: 24.5% or 48/196 contexts and Level 4: 19.4% or 34/175 contexts) verbs.

Table 4

Distribution of Level 3's and Level 4's Subjunctive Use in Task 3 according to Form Regularity

Form regularity	Level 3		Total	Level 4		Total
	#	%		#	%	
Form-specific	114	37.3	306	107	39.6	270
Irregular	83	32.8	253	80	34.6	231
Regular	48	24.5	196	34	19.4	175

Note. Level 3: X^2=8.903, *df*=2, Cramer's V=0.109, p=0.012; Level 4: X^2=20.292, *df*=2, Cramer's V=0.173, p<0.001.

Finally, the NSs' mood use on Tasks 1 and 3 differed (Table 5). On Task 1 (X^2=15.709, df=2, Cramer's V=0.107, p<0.001; Task 3: X^2=22.641), they used the subjunctive most often with regular verbs (55.9% or 219/392 contexts) and least often with form-specific irregular verbs (43.9% or 368/838 contexts), although the subjunctive use with irregular verbs was only slightly higher (45.0% or 67/149 contexts). On Task 3 (X^2=22.641, df=2, Cramer's V=0.200, p<0.001), subjunctive use was highest with irregular (71.6% or 139/194 contexts), followed by form-specific irregular (68.2% or 163/239 contexts), and, finally, regular (47.3% or 62/131 contexts) verbs.

Table 5

Distribution of NSs' Subjunctive Use in Tasks 1 and 3 according to Form Regularity

Form regularity	Task 1		Total	Task 3		Total
	#	%		#	%	
Form-specific	368	43.9	838	163	68.2	239
Irregular	67	45.0	149	139	71.6	194
Regular	219	55.9	392	62	47.3	131

Note. Task 1: X^2=15.709, df=2, Cramer's V=0.107, p<0.001; Task 3: X^2=22.641, df=2, Cramer's V=0.200, p<0.001.

In summary, on Task 1 NSs used the subjunctive most frequently with regular verbs and least frequently with form-specific irregular verbs, but the opposite trend was observed for Level 2 (i.e., subjunctive use was highest with form-specific irregular verbs and lowest with regular verbs). On Task 3, Levels 3 and 4 and the NSs produced the subjunctive least often with regular verbs, but NSs used the subjunctive most often with irregular verbs while Levels 3 and 4 used the subjunctive most often with form-specific irregular verbs. Furthermore, different patterns of subjunctive use across categories of form regularity were observed between Tasks 1 and 3 for the NSs.

3.2. Lexical item

The exploration of the variable lexical item starts with a type count of the lexical items. This count identifies the number of different verbs (e.g., *hablar*, 'to talk' and all of its forms (e.g., present subjunctive, preterit indicative) that appear in a dataset are grouped as one lexical item) produced by Levels 4 and 5 and the NSs individually. The NSs used 222 different lexical items in 2704 mood-choice contexts and each verb appeared an average of 12.2 times in the dataset. Level 5 used 175 lexical items in 2285-mood-choice contexts, with each lexical item appearing an average of 13.1 times in the dataset, and Level 4 produced 132 verbs in 2184 mood-choice contexts; each verb was produced an average of 16.5 times in the dataset. This result suggests that, in mood-choice contexts elicited from the three tasks under investigation, NSs have the largest vocabulary repertoire and Level 4 learners the smallest.[7] Next, a cross-tabulation shows the distribution of lexical items used in mood-choice contexts according to form regularity (Table 6). This analysis considers each lexical item only once. The results reveal that regular verbs are noticeably the most frequent (Level 4: 60.6% or 80/132 verbs, Level 5: 68.0% or 119/175 verbs, NSs: 70.7% or 157/222 verbs). Irregular verbs were produced more often (Level 4: 23.5% or 31/132 verbs, Level 5: 20.0% or 35/175 verbs, NSs: 18.0% or 40/222 verbs) than form-specific irregular verbs (Level 4: 15.9% or 21/132 verbs, Level 5: 12.0% or 21/175 verbs, NSs: 11.3% or 25/222 verbs). In other words, the largest vocabulary repertoire occurred with regular verbs and the smallest with form-specific irregular verbs.

---

[7] It is recognized that measuring lexical deployment is a complex and evolving issue (e.g., Jarvis, 2011) and that the current analysis is preliminary and exploratory.

Table 6

Distribution of Lexical Items Used in Mood-Choice Contexts according to Form Regularity

Form regularity	Level 4		Level 5		NSs	
	#	%	#	%	#	%
Irregular	31	23.5	35	20.0	40	18.0
Regular	80	60.6	119	68.0	157	70.7
Form-specific	21	15.9	21	12.0	25	11.3
Total	132	100	175	100	222	100

Following the examination of the lexical-item inventory and its relationship to form regularity, the frequency of use of individual lexical items was investigated. To begin to explore this issue, a global assessment of frequency of use was made. Firstly, the frequency of use of each lexical item for each participant group was identified (e.g., the NS group used forms of *estar* 'to be' 151 times). Secondly, after comparing the frequency trends that emerged from each dataset, three general groupings that characterize the data as a whole were identified. A collection of high-frequency lexical items appeared more than 100 times in each dataset (e.g., *ser* 'to be' and *poder* 'to be able to'), a collection of mid-frequency lexical items occurred on average one time for each participant (e.g., *estudiar,* 'to study' and *volver*, 'to return'), and a collection of low-frequency lexical items was produced on average less than one time per participant (e.g., *ayudar* 'to help' and *parecer* 'to seem'). The lexical items in the high- and mid-frequency groupings were the same for each participant group. The low-frequency grouping showed a range of the number of lexical items because each group produced a different number of lexical items (e.g., Level 4 produced 132 different lexical items while the NSs produce 222 different lexical items). Table 7 shows the results of this analysis and further divides the frequency groupings by the three categories of form regularity. Most lexical items, regardless of form-regularity category, were used infrequently (form-specific irregular: 70-147 lexical items, irregular: 14-18 lexical items, regular: 24-33 lexical items). The number of lexical items that fell into the mid-frequency grouping was relatively even across form-regularity categories (form-specific irregular: 5 lexical items, irregular: 6 lexical items, regular: 7 lexical items). The vast majority of high-frequency verbs were form-specific irregulars (5 lexical items). There was only one high-frequency irregular verb and no high-frequency regular verbs produced in mood-choice contexts by Levels 4 and 5 and the NSs. Further investigation into the mid-frequency grouping reveals a task effect; all of the lexical items that make up this grouping were verbs that were provided to the participants on Task 3. Thus, in general, the mid-frequency lexical items that were produced on average one time by each participant were also produced by the participants on Task 3. What is more, although each participant group produced a greater variety of low-frequency lexical items than the other groupings, high-frequency lexical items constituted almost half of the verbs produced in mood-choice contexts (Level 4: 48.0% or 1049/2184 contexts, Level 5: 47.6% or 1087/2285 contexts, NSs: 49.4% or 1336/2704 contexts).

Table 7

Distribution of Individual Lexical Items Used in Mood-Choice Contexts according to Frequency and Form Regularity

Form regularity	Low	Mid	High
Form-specific	70-147	5	5
Irregular	14-18	6	1
Regular	24-33	7	0
Total	108-198	18	6

Note. A range is provided in the low-frequency column because the three participant groups under investigation produced different quantities of lexical items.

Because the first two steps of the exploratory analysis of this dataset's lexical inventory shows that very few lexical items were used frequently, a detailed analysis of the role that the variable lexical item may have on explaining the effects of form regularity on mood use is limited to the high-frequency, form-specific irregular verbs. In other words, the initial assessment of the data revealed that the analysis employed to answer the second research question can only be conducted using a subset of the data. Thus, instead of investigating individual lexical items within all three categories of form regularity, the present study only makes observations about the relationship between the variable lexical item and mood use for high-frequency, form-specific irregular verbs. The final chi-square tests and cross-tabulations examine the relationship between mood use and high-frequency, form-specific irregular verbs, separately for Levels 4 and 5 and the NSs (Table 8). The lexical items included in this analysis were *haber* 'auxiliary verb 'has/had' and infinite of 'there is/are', *tener* 'to have', *estar* 'to be', *ser* 'to be', and *ir* 'to go'. The results indicate a significant relationship between mood use and lexical item for each participant group (Level 4: X^2=90.715, df=5, Cramer's V=0.294, p<0.001; Level 5: X^2=72.182, df=5, Cramer's V=0.258, p<0.001; NSs: X^2=63.290, df=5, Cramer's V=0.218, p<0.001.). Moreover, Level 4 and NSs showed the same scale of use across lexical items. Frequency of subjunctive use decreased in the following order: *haber > tener > estar > ser > ir*. Although each group used the subjunctive most often with *haber* and *tener*, the rate of subjunctive use for the remaining three verbs differed between Level 4 and the NSs on one hand and Level 5 on the other. The frequency of subjunctive use declined in the following order for Level 5 *haber > tener > ir > estar > ser*. The results also suggest that the learners varied their mood use across these five form-specific irregular verbs to a greater extent than the NSs. The difference in the rate of subjunctive use between the lexical item with the highest and lowest subjunctive use is the following: Level 4: range=36.5 percent, Level 5: range=33.4 percent and NSs: range=29.9 percent. This analysis demonstrates that, at least for high-frequency, form-specific irregular verbs, the rate of subjunctive use across lexical items is not uniform and that learners show greater differences in subjunctive use across lexical items than the NSs.

Table 8

Distribution of Subjunctive Use for High-Frequency, Form-Specific Irregular Verbs according to Form Regularity

Verb	Level 4		Level 5		NSs	
	#	%	#	%	#	%
Haber	50	43.1	137	61.2	171	64.8
Tener	35	18.2	101	54.0	96	49.0
Estar	20	15.6	44	44.0	62	41.1
Ser	20	8.2	72	27.8	140	39.2
Ir	14	6.6	90	47.1	74	34.9

Note. Level 4: X^2=91.139, df=4, Cramer's V=0.319, p<0.001; Level 5: X^2=60.302, df=4, Cramer's V=0.250, p<0.001; NSs: X^2=56.756, df=4, Cramer's V=0.219, p<0.001.

In summary, the exploration of the lexical inventory of mood-choice contexts demonstrated that most lexical items were regular verbs and were used infrequently. It also indicated that high-frequency verbs accounted for almost 50 percent of the data. A close analysis of mood use across high-frequency, form-specific irregular verbs revealed that the rate of subjunctive use was significantly different across lexical items. It showed that Level 4 and NSs were similar in their scale of subjunctive use across lexical item and that differences in the scale of use existed between these two groups and Level 5. It also demonstrated that NSs showed a smaller range of frequency of subjunctive use across lexical items than the learners (i.e., less variability).

4. Discussion

The current study sought to better understand the role that form regularity plays in the use of

verbal moods in Spanish by examining whether the variables task, lexical item, and participant group explain some of the varying effects found for form regularity on mood distinction in prior investigations. The findings are now discussed in relation to the three research questions.

The analysis of the variable task indicates that the effects of form regularity on mood use vary across data-elicitation tasks. Developmentally, these effects are observed with individual tasks before the dataset as a whole. Form regularity did not predict mood use for Levels 2 and 3 when all three tasks were analyzed together (Gudmestad, forthcoming), but there was a significant relationship between form regularity and mood use on Task 1 for Level 2 and Task 3 for Level 3. Interestingly, although form regularity predicted mood use for Levels 4 and 5 when all data were analyzed together, a significant relationship between mood use and form regularity was only found on Task 3 for Level 4 in the present analysis. Concerning the NSs, Gudmestad (2010) showed that form regularity predicted mood use when all of the data were analyzed together, but, in the current study, a significant relationship between mood use and form regularity was only found for Tasks 1 and 3. Furthermore, similarities and differences in mood use across the categories of form regularity were observed between analyses of all tasks together (Gudmestad, 2010, forthcoming) and individual data-elicitation instruments (the present study). In terms of similarities, Level 4 learners' subjunctive use was lowest with regular verbs when all of the data were analyzed together and on Task 3. NSs produced the subjunctive least frequently with form-specific irregular verbs on Task 1 and the dataset as a whole, and they used the subjunctive most often with irregular verbs on Task 3 and the dataset as a whole. In contrast, Level 4 learners produced the subjunctive most often with irregular verbs when the three tasks were analyzed together and with form-specific irregular verbs on Task 3. Other differences include the result that NSs used the subjunctive most often with regular verbs on Task 1 but with irregular verbs for the whole dataset and that their subjunctive use was least frequent with regular verbs on Task 3 and with form-specific irregular verbs when all tasks were analyzed together.

Moreover, there appears to be a similarity between the current group of NSs and the NS group investigated in Geeslin and Gudmestad (2008). Regular verbs predicted subjunctive use on an interview task in Geeslin and Gudmestad (2008) and in the current study subjunctive use was most frequent with regular verbs on Task 1. Both an interview and a monologic role play give speakers an opportunity for more spontaneous, extended oral production. The similarities between these tasks and the findings for the variable form regularity may be worthy of further investigation. Generally, the current study's findings seem to suggest that the inconsistent results in previous research (e.g., Collentine, 1997; Ross Veidmark & Umaña Aguiar, 1991) on the role that form regularity plays on mood distinction may in part be attributed to the diversity of tasks employed in the study of this issue. Thus, the present investigation underscores the importance of including a range of tasks to study a single phenomenon (Geeslin, 2010).

An assessment of the lexical inventory for Levels 4 and 5 and the NSs showed that a large number of individual verbs were used infrequently, which made a close analysis of the full range of lexical items difficult. This preliminary exploration could be helpful for moving forward this area of research since it showed that data need to be collected from tasks that elicit more instances of low- and mid-frequency verbs in order to carry out a more comprehensive study of this variable. Nevertheless, the exploration of the lexical inventory of the mood-use data appears to indicate that the variable lexical item helps to explain some of the inconsistent results found concerning the role that form regularity plays on mood distinction. A detailed examination of the relationship between the variables mood use and lexical item was only carried out on high-frequency, form-specific irregular verbs. The results on this subset of the data showed variability in subjunctive use among the five verbs for the three participant groups. These preliminary findings are in line with observations made by Poplack (2001), who found varying rates of subjunctive use across different verbs (i.e., lexical items) in her study of variable mood use in Canadian French. Additionally, the present findings appear to provide preliminary support for the hypothesis that vocabulary may play a role in explaining the effects of form regularity because different rates of subjunctive use were found for the five high-frequency, form-specific irregular verbs (cf. Zyzik & Gass, 2008). While the results for the NSs suggest that target-like subjunctive use varies according to individual lexical items, the range of frequency of subjunctive use for the Level 4 and 5 learners exceeded that of the NSs. This finding may indicate an intricate relationship between an increasing inventory of vocabulary (of verbs, in particular) and the ability to use the subjunctive with this growing lexical repertoire. Research on a wider range of proficiency levels and lexical items is needed to corroborate this observation. Thus, by offering a

careful characterization of the lexical repertoire produced in the dataset, the current study demonstrated that further investigation into the role of lexical item on mood use is warranted.

Finally, comparisons were also made across participant groups on the effects found for task and lexical item on mood use. Regarding the variable task and the examination of five proficiency levels and one group of NSs, the current study found more differences than similarities across participant groups. For example, the only similarity found across all participant groups was that no group showed a significant relationship between mood use and form regularity on Task 2. Perhaps the most notable contrast among Levels 4 and 5 and the NSs was the different range of frequency of subjunctive use across the high-frequency, form-specific irregular verbs. However, similarities were also found. For example, each of the three groups used the subjunctive most frequently with *haber* and *tener*. Despite some similarities, participant group seems to be another source of variation for the results found for the variable form regularity.

5. Conclusions and Future Directions

The present investigation has shown that the relationship between mood use and form regularity is complex and that the variables task, lexical item, and participant group seem to aid in explaining some of the differing results found across previous studies. The distribution of subjunctive use across categories of form regularity varied according to task. Within high-frequency, form-specific irregular verbs, learners and NSs used the subjunctive at varying rates, which suggests that lexical item may be another factor that predicts NS use and L2 development. Variation also existed across participant groups, both between L2 proficiency levels and between learners and NSs. In spite of this new information, several issues need to be investigated in future research in order to further promote our understanding of form regularity in particular and mood distinction in general. The source(s) of variation found across tasks, whether they are linguistic or extra-linguistic, need to be identified (cf. Geeslin, 2006). In other words, are the differences found across tasks for the variable form regularity due to characteristics of the data-elicitation instruments, the linguistic contexts that emerge from these tasks, or both? Data-elicitation tasks that allow low-, mid-, and high-frequency lexical items to be examined systematically should be designed. This way, a wider range of lexical items and their relationship to mood use and form regularity could be analyzed. In order to rigorously address the question of how vocabulary acquisition impacts the role that the lexicon plays on mood use and development, studies on the variable lexical item need to be conducted on a broader range of lexical items and L2 proficiency levels.

Appendix A: Additional Background Information on the Participants

Additional Participant Characteristics

| Group | # | Gender | | Age | | | Abroad experience | | |
		Female	Male	*M*	Range	#	*M*	Range	*SD*
Level 1	26	20	6	19.1	18-22	1	1 mo.	n/a	n/a
Level 2	35	29	6	19.0	18-21	1	1 mo.	n/a	n/a
Level 3	26	14	12	20.4	18-23	6	3 mos.	1-6 mos.	1.79
Level 4	23	17	6	21.5	20-24	16	3.88 mos.	1.5-10 mos.	2.71
Level 5	20	10	10	30.2	22-47	19	19.31 mos.	1.5 mos.-5 yrs. 1.5 mos.	16.51
NSs	20	10	10	30.2	25-42	20	4.84 yrs.	1.75-14.5 yrs.	2.84

Note. Abroad experience (study or work) for the learners refers to time in a Spanish-speaking country; for the NSs it refers to an English-speaking country.

Appendix B: Examples of Oral-elicitation Tasks

An English translation of an item from each data-elicitation task is provided. No Spanish version of the story segments of Tasks 2 and 3 is given because the stories were presented in English to eliminate other linguistic and comprehensibility factors that may have affected performance on the

instrument. The primary reason for this decision was to ensure that all learners understood the story. Tasks 2 and 3 did not require participants to provide extended oral production, so it would have been difficult to ascertain from the learners' language use whether they had understood the story if it had been written in Spanish. It was imperative that the participants understood the information being conveyed in these contexts, since two of the independent linguistic variables were coded at the level of discourse context.

Task 1, item 2

> *Eres un/a hijo/a único/a que tiene 10 años. Andas en el parque con tus papás y ves a una familia con 10 hijos. Empiezas a imaginar tu vida con muchos hermanos y se la describes a tus papás.*
>
> - *¿Serías el/la hijo/a menor, mayor o en el medio?*
> - *¿Qué te gustaría de esta situación familiar?*
> - *¿Qué te molestaría de esta situación familiar?*
> - *¿Qué querrías de tus papás?*
> - *¿Qué esperarían tus papás de ti?*
> - *¿Qué dudas tendrías respecto a esta situación?*
> - *¿De qué no dudarías?*
> - *¿Es importante hacer o no hacer ciertas cosas? ¿Por qué?*
> - *¿Cómo afectarían ciertas circunstancias específicas tu estado emocional (tus emociones)? Describe estas circunstancias y tus emociones.*
> - *Describe el ambiente en la casa.*
> - *¿Cómo cambiaría tu vida?*
> - *¿Habría algo que no fuera diferente?*

 'You are a 10-year old only child. You're walking in the park with your parents and you see a family with 10 children. You start to imagine what your current life would be like with many siblings and you describe it to your parents.

- Would you be the youngest, oldest, or middle child?
- What would you like about this family situation?
- What would bother you about this family situation?
- What would you want from your parents?
- What would your parents expect from you?
- What doubts would you have about this situation?
- What wouldn't you doubt?
- Is it important to do or not do certain things? Why?
- How would certain specific circumstances affect your emotion state (emotions)? Describe these circumstances and your emotions.
- Describe the environment in the house.
- How would your life change?
- Would there be something that wouldn't be different?'

Task 2, item 1
Clara, Pedro, and Tomás have just returned from semester break and are having dinner at a restaurant. They are catching up after being away from school for three weeks. Tomás had a great vacation. He and his family spent a week in Colorado and went skiing and hiking in the mountains. Talking about his trip, Tomás says:

"Me alegré de que _____."
'"I was happy that _____."'

Task 3, item 27
However, if they had gone to Italy, Javier could have had an opportunity see his grandparents. They have been living in Rome for many years. Ana says:

"Habría sido posible que se _____ (poner) en contacto con sus abuelos."
'"It would have been possible that _____ (to put) himself in contact with his grandparents."'

References

Bayley, Robert. (2002). The quantitative paradigm. In J. K. Chambers, Paul Trudgill, & Natalie Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 117-141). Malden, MA: Blackwell.

Berdan, Robert. (1996). Distentangling language acquisition from language variation. In Robert Bayley & Dennis Preston, (Eds.), *Second language acquisition and linguistic variation* (pp. 203-244). Amsterdam: John Benjamins.

Bosque, Ignacio. (Ed.) (1990). *Indicativo y subjuntivo.* Mardid: Taurus.

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

Ellis, Rod. (2008). *Study of second language acquisition*. Oxford: Oxford University Press.

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

Geeslin, Kimberly L. (2010). Beyond "naturalistic": On the role of task characteristics and the importance of multiple elicitation methods. *Studies in Hispanic and Lusophone Linguistics, 3*(2), 501-520.

Geeslin, Kimberly L., & Gudmestad, Aarnes. (2008). Comparing interview and written elicitation task in native and non-native data: Do speakers do what we think they do? In Joyce Bruhn de Garavito & Elena Valenzuala (Eds.), *Selected proceedings of the Hispanic Linguistics Symposium* (pp. 64-77). Somerville, MA: Cascadilla Proceedings Project.

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

Gudmestad, Aarnes. (2010). Moving beyond a sentence-level analysis in the study of variable mood use in Spanish. *Southwest Journal of Linguistics, 29*(1), 25-51.

Gudmestad, Aarnes. (forthcoming). Acquiring a variable structure: An interlanguage analysis of second-language mood use in Spanish. *Language Learning.*

Farley, Andrew P. (2004). Processing instruction and the Spanish subjunctive: is explicit information needed? In Bill VanPatten (Ed.), *Processing instruction: theory, research, and commentary* (pp. 227-39). London: Lawrence Erlbaum.

Jarvis, Scott. (2011, March). Constructs and dimensions of lexical deployment. Paper presented at the American Association for Applied Linguistics, Chicago, IL.

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

Montrul, Silvina. (2007). Interpreting mood distinctions in Spanish as a heritage language. In Kim Potowski & Richard Cameron (Eds.), *Spanish in Contact: Educational, linguistic, and social perspectives* (pp. 23-40). Amsterdam: John Benjamins.

Pérez-Leroux, Ana Teresa. (1998). The acquisition of mood selection in Spanish relative clauses. *Journal of Child Language, 23*(3), 585-604.

Poplack, Shana. (2001). Variability, frequency, and productivity in the irrealis of French. In Joan Bybee & Paul Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 405-428). Amerstdam: John Benjamins.

Ross Veidmark, Ronald, & Umaña Aguiar, Jeanina. (1991). La desaparición del subjuntivo español y sus implicaciones para el cambio lingüístico. *Filología y Lingüística, 17*(1-2), 193-202.

Schilling-Estes, Natalie. (2002). Investigating stylistic variation. In J. K. Chambers, Peter Trudgill & Natalie Schilling-Estes (Eds.), *The handbook of language variation and change*, (pp. 375-401). Malden, MA: Blackwell.

Scott, Dinorah. (2004). Mood variation in Spanish: A Neo-Gricean pragmatic explanation. (Doctoral dissertation, University of Illinois, Chicago, 2004). *Dissertation Abstracts International*, *64*(7).

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

Stokes, Jeffrey. (1988). Some factors in the acquisition of the present subjunctive in Spanish. *Hispania, 71*, 705-10.

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

Woolsey, Daniel. (2009). *Second language acquisition of the Spanish verb ESTAR with adjectives: An Exploration of contexts of comparison and immediate experience.* Munich: Lincom Europa.

Zyzik, Eve & Gass, Susan. (2008). Epilogue: A tale of two copulas. *Bilingualism: Language and Cognition, 11*(3), 383-385.

Selected Proceedings of the
14th Hispanic Linguistics Symposium

edited by Kimberly Geeslin
and Manuel Díaz-Campos

Cascadilla Proceedings Project   Somerville, MA   2012