Studies in Second Language Acquisition, 2010, **32**, 433–463.
doi:10.1017/S0272263110000033

AN EXPLORATION OF THE RANGE AND FREQUENCY OF OCCURRENCE OF FORMS IN POTENTIALLY VARIABLE STRUCTURES IN SECOND-LANGUAGE SPANISH

Kimberly L. Geeslin
Indiana University

Aarnes Gudmestad
Virginia Polytechnic Institute and State University

This article adds to the growing body of research focused on second-language (L2) variation and constitutes the first large-scale study of the production of potentially variable grammatical structures in Spanish by English-speaking learners. The overarching goal of the project is to assess the range of forms used and the degree to which native and L2 speakers of Spanish differ in several independently defined syntactic or discourse-based contexts. The contexts examined in the current study have been the object of sociolinguistic research in monolingual environments and include the following: copula contrast, mood distinction, past-time reference, future-time reference, and subject expression. Interview data from 16 English-speaking learners and 16 native speakers of Spanish from a variety of countries, all of whom are part of a single speech community in the United States, are examined. The analysis focuses on the range of forms used in each of the contexts investigated and the frequency with which these forms

The authors are engaged in a continuing collaboration in which the order of names alternates from one paper to the next. We would like to thank Leanna House for her help with the quantitative analysis and her insightful comments regarding the practices and assumptions in statistics. All errors are our own.

Address correspondence to: Kimberly L. Geeslin, Department of Spanish & Portuguese, Ballantine Hall 844, Indiana University, Bloomington, IN 47405; e-mail: kgeeslin@indiana.edu.

© Cambridge University Press, 2010 0272-2631/10 $15.00

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

appear. A possible relation of individual characteristics, such as country of origin, years of language study, and time spent abroad, to this frequency of use is also considered.

Inherent in most definitions of communicative competence is a second-language (L2) learner's ability to use variants of a language in all areas of the grammar (e.g., phonology, syntax) in context-appropriate ways (e.g., Canale & Swain, 1980). The acquisition of such variants and the rules that govern them allow learners to convey important personal, social, and pragmatic information. Additionally, an understanding of variation in the L2 plays an essential role in the development of theories of SLA. Much progress has been made over the past decade in the degree to which psycholinguistic models have developed plausible accounts of the unconscious figuring of the relative effects of social and linguistic factors on language use and in the quantity of empirical evidence available to document the fact that L2 input, input processing, and L2 language use are socially mediated (see Preston, 2000; Tarone, 2007). Although it is generally recognized that learners vary their language use in the process of acquiring a given structure and that learners can acquire the ability to vary their language according to social norms, research on the SLA of variable structures is not widespread. This paucity of research may be due to the lack of information available in the L2 literature with regard to the contexts in which variation occurs and to the cross-disciplinary nature of this research. The goal of the present study is to respect practices and understandings across fields while advancing knowledge of L2 variation and theories of SLA. Potentially variable structures are examined, with careful attention to recent developments in the way contexts for variation are identified, while maintaining a view of interlanguage as a complex system in which multiple elements may interact and influence language use.

The term *potentially variable structures* stems from long-held assumptions (and debate) in the field of sociolinguistics and from recent developments in work on language use in Spanish. From the inception of modern sociolinguistic research, a linguistic variable has been understood to entail two or more ways of conveying the same referential meaning or truth value (Labov, 1972). Thus, phonological variables were considered to be the most appropriate features of the grammar for inclusion in variationist studies, and the extension of such research beyond the sound system, although not uncommon (e.g., Sankoff & Vincent, 1977; Weiner & Labov, 1983), has been the object of considerable debate (e.g., Cheshire, 1987; Labov, 1978; Romaine, 1984; Winford, 1996). Despite this controversy, there are reasonable arguments for applying the same predictive studies to nonphonological variables in order to examine the multiple influences

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

on (non)use of a variant. For example, Cheshire ordered linguistic variables along a continuum whereby meaning distinctions increased from morphophonemic to syntactic variables and argued that many cases of morphological variation can also be analyzed through existing methods (i.e., multivariate analyses). An even more important point made by Cheshire is that although instances of syntactic and semantic variation may not constitute strictly defined linguistic variables, this does not mean that such variation does not exist or that it should not be studied. Instead, she argued that spontaneous spoken language ought to be analyzed and that this may require the development of new methods of analysis. There is certainly value in examining variation in nonphonological variables in the domain of both the first language (L1) and the L2, and there is a direct connection among the arguments made by Cheshire, recent work in L1 and L2 Spanish on the five structures analyzed here, and the goal of developing the methods necessary to continue this line of research.

There are two aspects of this debate that pose challenges for research on certain grammatical structures in Spanish (as well as in other languages). First, it is expected that the analysis of the predictors of a given dependent variable must be limited to contexts in which the semantic meaning conveyed by the two (or more) variants is the same. This equivalence in meaning is more difficult to demonstrate with syntactic and semantic variation because the discourse features (which can be syntactic, semantic, or pragmatic) generally associated with the meaning conveyed tend to overlap, operate in unison, and depend on the intended meaning of the speaker. Thus, these variants are usually associated with a given meaning, but use is not categorical. For example, although subject expression may be related to the person and number of the verb, it is simultaneously influenced by factors such as switch reference (i.e., whether the subject of the preceding finite verb is the same as the subject of the current token) and potential ambiguity as well as tense, mood, and aspect of the verb form (Bayley & Pease-Álvarez, 1997). For the grammatical structures examined in the current study, there is no single feature (i.e., meaning) that predicts the categorical use of a form. Another challenge is to objectively identify the contexts in which the variable occurs (i.e., to describe the context without reference to the form itself). These challenges do not mean that existing definitions are inappropriate but rather that new methods must be developed.[1]

Before continuing with examples of cases that cannot be dealt with using established methods, it should be acknowledged that these conditions can be met for some nonphonological structures. For example, in Young's (1991) study of the linguistic and nonlinguistic predictors of plural marking in L2 English, the context was defined as all conceptually plural noun phrases and the variants considered were the presence or absence of plural markers (e.g., *–s*). This study is an example of the

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

appropriate application of variationist methodology beyond the sound system and shows that the contexts in which the variable occurs can in some cases be defined independently of the variants under examination. Other examples include studies on the use of *–ing* (Adamson & Regan, 1991) and past-tense marking (Bayley, 1996; Wolfram, 1985) in L2 English and *ne* deletion in the *ne…pas* "not" structure (Regan, 1996) and the impersonal pronouns *nous* "we" and *on* "one" (Rehner, Mougeon, & Nadasdi, 2003) in L2 French.

In contrast to Young (1991), the current study lays the groundwork for five potentially variable structures whose semantic meaning and context of occurrence require careful consideration. An excerpt from spoken corpus data provided in (1) and described in Díaz-Campos and Geeslin (in press), in which the syntactic structure copula plus adjective occurs three times with the adjective *feo* "ugly" and in which *ser* "to be" alternates with *estar* "to be," illustrates this need for further examination.

(1) *Nosotros hemos conseguido parejas besándose ahí, metiéndose las manos,* ***eso es muy feo****... "Mira, eso que tú estás haciendo,* ***eso está feo*** *¿No te da pena eso? Busca un mejor sitio para eso, pero ¿en la calle? ¿ah? ¡****Eso es feísimo****, eso no se puede!"*
"We have found couples kissing over there, caressing each other with their hands, **that is (*ser*) ugly**... 'Look, what you are doing, **that is (*estar*) ugly**. Aren't you ashamed? You need to look for a better place, but in the street? Ah? **That is (*ser*) very ugly**, that cannot be possible!'"

In their study of Venezuelans from four different generational groups, Díaz-Campos and Geeslin argued that such use must be described in terms of multiple independent variables. The contexts examined were defined syntactically (i.e., copula plus adjective) and limited to occurrences of *ser* "to be" and *estar* "to be." This is in fact the approach taken by nearly all studies of copula contrast (e.g., Geeslin & Guijarro-Fuentes, 2008); nevertheless, two issues arise. First, within the copula-plus-adjective context, the meaning conveyed is the same only if it is broadly defined as attributive. It has been argued that the two copular verbs in question actually convey different meanings (Fernández Leborans, 1999), which are often defined as a composite of several different pragmatic and semantic features. It is clear that these two verbs do not meet the traditional definition of variants of the same linguistic variable. It is likely that within the copula-plus-adjective context, there are some obligatory contexts and some variable ones, in which the same meaning is conveyed; however, objective means for distinguishing such contexts from one another do not yet exist. Second, the full range of forms that can occur in the copula-plus-adjective context is not represented in the literature, because other verbs such as *parecer* "to seem," *sentir* "to feel," and *ponerse* "to become" also occur with some regularity in this

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

context. Even studies that initially included a range of forms in the analysis, such as Geeslin (2005), collapsed them all into a non-*estar* category for the purpose of statistical analysis. Although the statistical analysis of multiple independent variables employed in this study required that the forms be considered together, it will be argued that overlooking the full range of forms that occurs in a given context is in direct conflict with the conceptualization of interlanguage as a complex system in which the many forms used to fulfill a given function in the developing grammar are interrelated. In other words, the full range of forms will provide important details about the differences between L1 and L2 Spanish. The interdisciplinary nature of the current study therefore becomes clear: There is both a need in sociolinguistics to address variation that departs from traditional assumptions about linguistic variants as well as a requirement in L2 research to treat interlanguage as an independent system of interrelated components. Addressing these two concerns will improve research in both fields.

Both the contexts of occurrence of these five potentially variable structures and the full range of forms that occur in these contexts should be examined; furthermore, the range of forms should be considered before the contexts of occurrence. This is similar to the approach taken by Cheshire (1987): By first examining the full range of forms that occur in objectively defined contexts, the current study is not confined by the traditional definition of a linguistic variable and, thus, new approaches can be developed. Only once it is has been determined which forms occur in a given context, which initially should be broadly defined, can the analysis be focused on a subset of those forms in a more narrowly defined context (i.e., one in which only the same meaning is conveyed by two or more variants). For each of the grammatical structures examined here, there is either a broader range of forms that occur in a given context than has been previously examined in a single analysis or the context under examination must be more broadly defined than has been done previously. The use of the term potentially variable structures reflects the possibility that the forms that occur in these structures, or a subset of these forms, exhibit variation; however, the necessary analyses to demonstrate this do not yet exist. Before these potentially variable structures can be the object of a multifactor variation analysis, the questions raised here of both the meaning conveyed and the range of forms in variation must be addressed. This may serve as a first step that leads to a more traditional multivariate analysis or to other types of analyses. The current study addresses the issue of the range of forms that occur in contexts previously analyzed as if a binary dependent variable existed for these structures; this is a necessary first step in understanding the variation that may occur in these contexts.

Although the primary goal of this investigation—namely, to demonstrate the need for analyses that focus on contexts that are broadly

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

and independently defined and the full range of forms that occur in those contexts—is methodological, new information about the differences between learners and native speakers (NSs) of Spanish is also provided. Examining the range of forms that occur in a given context and the frequency with which they occur provides a new means for comparing these groups. The analysis of the relationship between this frequency of occurrence and the individual characteristics of the speakers provides important insight into the way in which individual experiences affect acquisition. Because the data come from advanced learners, rather than from learners at many levels, no attempt will be made to describe the path of acquisition of these structures. Instead, a concrete, detailed analysis of L2 Spanish will be provided and this interlanguage grammar will be compared to the grammar of NSs. The focus on advanced learners will also impact the development of theories of SLA, because the current study examines structures that are acquired late in the process of acquisition and that have typically not been examined among populations that possess the grammatical resources necessary to demonstrate socially mediated variation. In other words, the development of theories of SLA that incorporate variation requires a detailed understanding of how learners use language even beyond the first stages of development. However, these learners are not at the end state of development; they are quite advanced, and those who pursue further study will likely develop more nativelike use. Thus, the current study documents the range and frequency of forms used to perform a single morphosyntactic function by a group of learners and by a comparable NS group, and this analysis of language use stands to impact the understanding of L2 use, the process of SLA, and the variability inherent in both.

In sum, the current study explores the range and frequency of forms used in five different potentially variable structures to explore potential dependent variables for studies of variation in L2 Spanish. These forms and their frequency of occurrence provide an objective means to compare English-speaking learners to a similar group of NSs and to address the degree to which individual characteristics, such as language-learning experience, are linked to variability. This investigation constitutes a first step in research on these potentially variable structures and provides the basic information needed to pursue further analyses of each of these structures.

BACKGROUND

Before reviewing the five potentially variable structures analyzed in the current study, an overview of the research on L2 variation will be provided.

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

SLA of Variation

Second-language variation has been described in depth and shown to exist within and between speakers and across speech contexts and learning tasks (e.g., Bayley & Langman, 2004; Ellis, 1999; Young, 1991). Work by Tarone (e.g., 2007) and Preston (e.g., 2000) has called for theories of SLA that incorporate L2 variation and recognize the variable nature of input. This field has gained momentum by employing variationist analyses of the relevant social and linguistic predictors of a given variant for a group of learners (e.g., Adamson & Regan, 1991; Rehner et al., 2003), and studies on variation in L2 Spanish have emerged (e.g., Geeslin & Guijarro-Fuentes, 2006; Gudmestad, 2006).

One key development in these studies is the distinction made between variation that exists only in learner language and variation that exists in both the learner language and the target language. Variation exclusive to learner language has been called vertical variation (Adamson & Regan, 1991) or Type I variation (Rehner, 2002). Although the current study examines structures that vary in the target language (horizontal or Type II variation), there are also developmental factors to be considered. In fact, although the grammatical structures examined may vary for NSs and nonnative speakers (NNSs), it may be the case that obligatory contexts (those in which no variation occurs) for the NSs are also variable for the NNSs (Geeslin & Guijarro-Fuentes, 2006). Cameron (1994) called the contexts in which variation may occur the envelope of variation, but it is possible that this envelope may be different for both groups considered here. This possibility further supports the idea that research on the nonphonological potentially variable structures in the current study may require new or additional methods of analysis. Thus, the current study defines contexts independently of the forms that occur in them, to avoid the exclusion of variable contexts in the learner data as a result of assumptions about NS use.

A second important fact about recent research on L2 variation is that the analyses used have generally included social and linguistic predictors of a variant and have sought to predict the presence or absence of that variant against all other possibilities. This approach works well for Type I variation: The presence of the targetlike variant is compared to all others (Young, 1991). This means of analysis is also effective for some Type II variation, such as the comparison of a newer form against the older expected form, as in the case of French *on* "one" in contexts in which *we* is the understood subject of the verb (Rehner et al., 2003).[2] Nevertheless, the approach to the examination of the structures considered in the current study differs in that all forms that occur in a discourse or syntactically defined context are viewed as potential categories of a dependent variable. This is a necessary development for research on the structures

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

examined here, especially in a L2 context, where there is even less certainty about the so-called envelope of variation. For example, although NSs of Spanish may agree on the contexts in which full lexical subject pronouns are required, this consensus cannot be assumed for learners. In the current study, therefore, an analysis of the range and frequency of use of all forms in a given context will precede an analysis of the predictors of a single form. Thus, it is not assumed that all of the potentially variable structures are linguistic variables, even though this provides a means for learning more about language use in those contexts. The eventual goal for research on this topic will likely be a predictive analysis in which the linguistic and social factors that constrain such variation are identified; however, the current study may reveal that this type of analysis is not appropriate for one or more of the structures examined.

Variable Structures in Spanish

Subject Expression. Spanish allows the subject of an inflected verb to be variably explicit or understood as a function of context without being overtly expressed. As illustrated in (2), when introducing a referent for the first time, as in (2a), the subject may be expressed as a full noun phrase. Subsequent mentions of that referent may be null pronouns, as in (2b) and (2c), or overt subject personal pronouns, as in (2d), when contrast is implied or the referent is ambiguous.

(2) (a) ***Juan*** *es mi amigo de Venezuela*. (b) ***ø*** *Tiene tres hijos y* (c) ***ø*** *trabaja con Maria*. *Julia dice que* (d) ***él*** *es muy inteligente*.
"**John** is my friend from Venezuela. [**He**] has 3 kids and [**he**] works with Mary. Julie says that **he** is very intelligent."

The null subject is expected in cases in which a referent is clear and no contrast is implied and is obligatory with certain verbs, such as those that express atmospheric conditions (e.g., *llover* "to rain"), and in certain syntactic structures, such as within a relative clause.

Sociolinguistic research on subject expression addresses the conditions that favor the use of overt subjects (e.g., Bayley & Pease-Álvarez, 1997; Silva-Corvalán, 1982). In general, such studies have focused on subject personal pronouns (i.e., the human subject of a tensed verb) in variable contexts (Otheguy & Zentella, 2007). In addition to extending research to other Spanish-speaking populations and confirming the effect of the independent variables used in earlier work, studies on subject expression have explored the relationship between overt subjects and /s/-weakening (e.g., Cameron, 1994; Cameron & Flores-Ferrán, 2004) and how language contact influences subject expression (e.g., Otheguy & Zentella).

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

Research on subject expression in L2 Spanish has been informed primarily by generative and optimality theory (e.g., Isabelli, 2004; Lafond, Hayes, & Bhat, 2000). There are no L2 studies that have addressed the frequency of occurrence or the range of contexts in which overt subjects occur (but see Lozano, 2002, and López-Ortega, 2003, for other related studies). Moreover, sociolinguistics has largely been limited to a subset of verbal subjects: those with human referents, where a pronoun (as opposed to a full noun phrase) would be expected and where variation is deemed to have the possibility of occurring, even though variable contexts have yet to be exhaustively or empirically identified (Otheguy & Zentella, 2007). Thus, no research to date has examined the range of forms (null, full lexical noun phrases, and pronouns) in all inflected verbal contexts for NSs or learners. Because the definition of interlanguage relies on an analysis of a whole system separate from a comparison to the target, this type of broad analysis is a logical starting point that will lead to concrete information about the range of forms used in other contexts and the frequency with which these forms occur.

Copula Use. In Spanish, *ser* "to be" and *estar* "to be" occur with adjectives. In some cases, only one co-occurs with an adjective (e.g., *descalzo* "barefoot" with *estar* "to be"), but 80% of Spanish adjectives, as in (3), allow alternation depending on the meaning expressed (Mesa Alonso, Dominguez Herrera, Padrón Sánchez, & Morales Aguilera, 1993).

(3) a. *Mario **está** muy guapo.*
"Mario **looks** very handsome (today)."
b. *Mario **es** muy guapo.*
"Mario **is** very handsome."

Additionally, sometimes verbs such as *parecer* "to seem" and *sentir* "to feel" co-occur with adjectives. Although theorists disagree, the contrast between the copulas and the resulting meaning changes are said to be linked to features such as the type of predicate (Fernández Leborans, 1999), the frame of reference (Clements, 2005), and the class of adjective (Vañó-Cerdá, 1982). It is widely—although not unanimously—agreed that copula contrasts result from the intention of the speaker to highlight various meanings and, thus, several features of the discourse context are simultaneously implicated in this contrast (Clements). This is precisely why variable contexts are difficult to identify and new methodology is in order.

Sociolinguistic research on the copula contrast began with interest in Mexican American Spanish (Silva-Corvalán, 1994b), where the extension of *estar* "to be" to new contexts was linked to a loss of linguistic constraints (i.e., circumstantiality, frame of reference, and susceptibility to

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

change) over time. Such variability has since been demonstrated for dozens of populations and linked to linguistic factors, such as predicate type, changeability of the referent, experience with the referent, adjective class, frame of reference, and animacy of the referent, as well as to level of education, access to the written standard, social class, gender, elicitation task, and knowledge of a regional language (Cortés-Torres, 2004; Geeslin & Guijarro-Fuentes, 2008; Gutiérrez, 2003).

Early work on L2 copulas identified stages of acquisition for all functions of the copular verbs (e.g., progressive, locative) and demonstrated a relatively consistent path of acquisition (Ryan & Lafford, 1992; VanPatten, 1987). Conflicting results were found in the order of acquisition of *estar* "to be" with adjectives, which is precisely where NS variation exists. In response to these inconclusive results, studies have examined the linguistic and social correlates of copula selection for several learner populations in the copula-plus-adjective structure (Geeslin, 2003; Geeslin & Guijarro-Fuentes, 2006). In general, the same linguistic predictors of NS copula choice can also be applied to NNS use, and the shifting importance of these predictors over time is an effective way to describe the process of SLA (Geeslin). In contrast, social characteristics are less consistent predictors of L2 copula choice (Geeslin & Guijarro-Fuentes). Although variability in L2 copula choice has been documented, no study has examined the range of forms used in the copula-plus-adjective structure and compared frequencies of occurrence between advanced NNSs and NSs.

Mood Use. Spanish makes mood distinctions primarily in subordinate clauses, and it is often explained that semantic triggers in the matrix clause influence mood use (Butt & Benjamin, 2004).[3] The indicative traditionally appears in contexts such as those of assertion, definiteness, and habitual actions, whereas the subjunctive is used in contexts such as those of volition, comment or emotion, uncertainty, and indefiniteness and in adverbial clauses of futurity. Examples of mood use in contexts of assertion and volition are provided in (4) and (5), respectively.

(4) *Sé que la fiesta* ***es*** *el sábado.*
"I know that the party is on Saturday."
(5) *Quiero que me* ***ayudes*** *con los quehaceres.*
"I want that you help me with the chores."

The past subjunctive can also occur with *si* "if" clauses for hypothetical or nonexistent situations. In independent clauses, mood distinctions are also made after expressions such as *ojalá* "I hope," *quízas* "maybe," and *tal vez* "perhaps," and the past subjunctive can be used to mitigate speech.

Although mood distinction has been associated with speaker intent and assertion (e.g., Haverkate, 2002), sociolinguistic investigations on Spanish in Latin America, Spain, and the United States have largely carried

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

out sentence-level analyses (e.g., Blas Arroyo & Porcar Miralles, 1997; Lynch, 2000; Murillo Medrano, 1999). These studies, which focused on dependent clauses (for an exception, see Bayerová, 1994), have shown that speakers rarely respond categorically to the same linguistic contexts. Moreover, the rate of use and selection of verbal moods varies across semantic contexts. For example, the use and selection of the subjunctive is high after matrix clauses of volition, lower after matrix clauses of uncertainty and comment, and even less frequent after matrix clauses of assertion. Analyses of apparent time data have also suggested that there is a change in progress toward an increased use of the indicative (e.g., Silva-Corvalán, 1994a). Within the linguistic contexts investigated to date, mood distinction is clearly variable among NSs. However, because of the focus on traditionally defined obligatory and variable contexts in subordinate clauses (e.g., Lynch; Silva-Corvalán), the contexts in which the subjunctive and indicative moods are in alternation have not been empirically defined. In the current study, therefore, the analysis of mood use includes all finite-verb contexts.

Research on mood distinction in L2 Spanish has examined issues such as form regularity and comprehension and shown that this distinction is acquired late (see Collentine, 2003). Despite the breadth of research, L2 studies have only recently incorporated insights from sociolinguistics by moving away from error analysis to a study of frequency and predictors of NS and NNS use (e.g., Geeslin & Gudmestad, 2008; Gudmestad, 2006, 2008). For example, in an analysis of cross-sectional data, Gudmestad (2008) showed that the semantic category conditioned mood use earlier than the discourse-level variables of time reference and hypotheticality and that the effect of these variables preceded that of form regularity. Additionally, the most advanced learners approached NS norms of frequency and predictors of use. Much of this research has been limited to traditional subjunctive contexts (for exceptions, see Geeslin & Gudmestad; Gudmestad, 2008), even though it is known that NNSs vary in contexts where NSs do not. This observation further motivates the analysis of mood distinction in all finite-verb contexts.

Past-Time Reference. In contexts of past-time reference, as indicated by lexical temporal indicators (e.g., *ayer* "yesterday") or by the sentence- or discourse-level context, Spanish marks aspect and relevance to the present (Butt & Benjamin, 2004). Examples (6) and (7) demonstrate the distinction between perfective and imperfective aspects.

(6) *Juan **desayunaba** cuando sonó el teléfono.*
"Juan **was eating** breakfast when the phone rang."
(7) *Juan **desayunó** antes de ir a la escuela.*
"Juan **ate** breakfast before going to school."

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

In (6), the verb *desayunar* "to eat breakfast" is marked with imperfective aspect to indicate an ongoing action (interrupted by another event), and in (7), perfective aspect indicates an action completed prior to the subsequent event (going to school). In perfective contexts, the present perfect denotes events that have current relevance (Brugger, 2001), as in (8).

(8) *No **he desayunado** y tengo hambre.*
"I **have not eaten** and I am hungry."

There is an absence of empirical documentation that the contrast between the imperfect and one of the perfective forms (i.e., preterit, present perfect, pluperfect) is variable. However, the distinction between the preterit and the present perfect has been linked to several factors. More frequent use of the present perfect is associated with Spain, where the use of this form is increasing in contexts that are temporally indeterminate (Schwenter & Torres Cacoullos, 2008). Similarly, the present perfect has been linked to urban centers, current relevance or recentness, the presence of temporal adverbs, and language contact (e.g., Brugger; Klee & Ocampo, 1995; Mackenzie, 1995; Schwenter & Torres Cacoullos).

Second-language acquisition research has focused nearly exclusively on the distinction between the preterit and the imperfect with the goal of charting the emergence of past-tense marking (Bardovi-Harlig, 2000; Salaberry & Shirai, 2002). Although English has been the primary focus of investigation, L2 Spanish is well studied (see Salaberry, 2008). Two theories of development, the aspect hypothesis (e.g., Andersen & Shirai, 1994) and the discourse hypothesis (Bardovi-Harlig), have emerged from this research and are currently viewed as equally valuable in describing the path of acquisition of past-tense forms. Both hypotheses are supported by analyses of L2 Spanish (Andersen, 1991; Cadierno, 2000; Liskin-Gasparro, 2000). Despite this wealth of research, no study that addresses the full range of forms that appear in contexts of past-time reference and the frequency with which these forms occur for relatively advanced learners (i.e., not when these structures first emerge) has been conducted. Of all the structures included in the current study, contexts of past-time reference may be those in which variation is least expected, but it is possible that some contexts, even defined according to semantic meaning, are indeed variable.

Future-Time Reference. In Spanish, contexts of future-time reference, as indicated by lexical temporal indicators (e.g., *mañana* "tomorrow") or by context, can be marked verbally with the morphological (synthetic) or the periphrastic (analytic) future, as in (9a) and (9b), respectively. It

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

is also possible to mark future time with the present indicative and with the subjunctive in subordinate clauses with adverbial conjunctions of time (Butt & Benjamin, 2004).

(9) a. ***Compraré*** *un coche*.
"I **will buy** a car."
b. ***Voy a comprar*** *un coche*.
"I **am going to buy** a car."

Sociolinguistic research on data from the Southwest of the United States, Mexico City, and Morelia, Mexico (Gutiérrez, 1995), Barranquilla and Colombia (Orozco, 2005), Puerto Rico, Venezuela, the Dominican Republic, and Chile (Silva-Corvalán & Terrell, 1989) has shown that NSs of Spanish use the morphological future, the periphrastic future, and the present indicative with varying frequencies across dialects. Although in most cases the periphrastic future is more frequent than the morphological future, the opposite is true in Peninsular varieties (e.g., Blas Arroyo, 2008). The factors found thus far to condition the forms in future-time contexts include certainty of the event, temporal distance, grammatical person and number, negation, presence of a lexical temporal indicator, clause type, and a speaker's age and gender.

The SLA of verb forms for future-time reference has been investigated for English (e.g., Bardovi-Harlig, 2004) and French (e.g., Nadasdi, Mougeon, & Rehner, 2003), but research on Spanish has been largely limited to incidental acquisition (Lee, 2002; Rossomondo, 2007). There appears to be no study on L2 Spanish that examines the use of verb forms in future-time reference. One interesting contrast between L2 English and L2 French is that whereas learners of English use the morphological future more frequently, learners of French use the periphrastic future more often (e.g., Bardovi-Harlig; Nadasdi et al.). Thus, the present investigation fills a void in the current literature and has implications for researchers who work on other languages.

THE CURRENT STUDY

The present study is a large-scale examination of five potentially variable structures in Spanish and the forms that NSs and NNSs use in those contexts. One unique feature of this project is that the contexts are independently defined (i.e., no reference to the forms that occur is made) and the full range of forms used in those contexts is analyzed. The individual characteristics associated with variability in the learner group were also examined. It is only after research of this type is

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

conducted that a closer examination of the predictors of any one of those forms becomes meaningful. The current study was guided by three research questions:

1. In a sociolinguistic interview, what range of forms do language learners use to express potentially variable structures in Spanish and with what frequency do they use these forms?
2. How does this language use compare to that of NSs on the same task?
3. Does the production of these forms vary across groups according to gender, time abroad, and years of language study (i.e., independent nonlinguistic variables)?

Participants

The first participant group included 16 NSs from Argentina (two males), Chile (one female), Colombia (two females and two males), Ecuador (one female), Mexico (one male), Spain (two females and two males), Puerto Rico (one female), the southwestern United States (one male), and Uruguay (one female). All were residing in the United States and ranged in age from 24 to 37 (M = 29.8 years). On a 25-item proficiency test, their scores ranged from 22 to 25 (M = 23.6). The second group consisted of 16 English-speaking learners of Spanish (10 male, 6 female) between the ages of 22 and 35 (M = 26.4 years), whose proficiency test scores ranged from 16 to 25 (M = 20.8). They had spent between 3 months and 11 years (M = 18.1 months) in a Spanish-speaking country (Argentina, Chile, Colombia, Costa Rica, Ecuador, Mexico, Nicaragua, and Spain). When the range and mean are recalculated to exclude one speaker who spent 11 years abroad, the range is 3–27 months (M = 11.2 months). The NNSs had studied Spanish formally for 3–17 years (M = 8.07 years).

The current study includes a group of NSs because a complete picture of how NSs use the five potentially variable structures under investigation does not yet exist. Moreover, the range of their countries of origin reflects the diverse input the NNSs have received. All participants were graduate students, were current or former instructors of undergraduate Spanish, and interacted socially. Finally, both groups completed the same task and, thus, had the same opportunities for production in response to the same topics of conversation.[4] Although dialect-specific variation or monolingual Spanish cannot be addressed here, the NSs' diversity of background and participation in the same speech community make them appropriate for comparison to the NNSs.

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

Data Elicitation Instruments

All participants completed a background questionnaire, a language proficiency test, and a sociolinguistic interview. The background questionnaire was a short fill-in instrument that provided the biographical information about the participants. The language proficiency test was a 25-item multiple-choice activity contextualized in a story in Spanish. It covered various grammatical concepts, such as object pronouns and verb forms, and participants selected a response from three options. This instrument provided an objective means for comparing participant groups that was not limited to the structures analyzed in the current study. Even though the L2 participants are considered to be near-native based on program-specific employment requirements, this additional measure is included to demonstrate their range of abilities, and the term *advanced* (rather than near-native) is used based on that variability.

The third task was a 30-min sociolinguistic interview with two NSs of Spanish (one female and one male, each from a different country). This interview was conducted in a quiet language laboratory, and the interviewers were acquaintances of the participants. The interview questions were the same for all speakers, but conversations varied according to the interests and level of detail provided by each participant. The interview questions covered topics such as plans for the future, recent and past experiences, and opinions about various political and social issues. For example, participants were asked about their plans for the upcoming weekend, their career goals for the next 5 years, and their reaction to the recent elections. The interviews were digitally recorded and transcribed.

Data Coding and Analysis

The contexts in which each structure occurs were described independently of the expected forms, and each structure was treated as a distinct dependent variable that could be influenced by any of the nonlinguistic independent variables in the study. For subject expression, each subject position of a tensed verb was an individual token. Each tensed verb form was coded for the mood used. This broad analysis of mood distinction contexts is based on an objective syntactic definition and includes subjunctive forms used in contexts in which they do not traditionally occur. All contexts that had the syntactic structure copula plus adjective were coded for copula selection. In the current study, copula is defined as a verb that connects a referent with a predicate adjective, and, along with *ser* "to be" and *estar* "to be,"

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

verbs such as *sentir* "to feel" and *parecer* "to seem" are also included in this definition. All contexts where reference was made to the past or future (at the sentence or discourse level) were coded for the form of each finite verb. Each grammatical structure, the means of identification, and the range of forms that appeared in that context are summarized in Table 1. Nontargetlike forms were rare for the NNSs, but those that did occur were coded along with targetlike forms, provided they were similar enough to nativelike forms to be identified. Any questionable cases were discarded or discussed until the authors agreed on their classification. The data were also coded for L1 (Spanish vs. English), gender, time abroad (NNSs only), and years of Spanish-language study (NNSs only).

To answer the first two research questions, cross-tabulations were carried out for all categories of each dependent variable and each participant group. The categories of each dependent variable and the rate at which they were used were compared between speaker groups. To answer the third research question, cross-tabulations were conducted for gender for both speaker groups separately and for time abroad and length of language study for the NNSs only.[5]

Table 1. Coding for the dependent variables

Dependent variable	Definition of a token	Categories of the variable
Subject expression	The subject of all finite verb forms	Null, lexical noun phrase, personal pronoun, demonstrative pronoun, indefinite pronoun, interrogative pronoun
Copula use	All copula-plus-adjective structures	*Ser* "to be," *estar* "to be," *parecer* "to seem," *sentir* "to feel"
Mood use	The mood chosen for all finite verbs	Subjunctive, nonsubjunctive
Past-time reference	All verb forms in contexts that refer to the past (based on features of the sentence or the surrounding discourse)	Present perfect, preterit, imperfect, past subjunctive, pluperfect, present, conditional, present subjunctive
Future-time reference	All verb forms in contexts that refer to the future (based on features of the sentence or the surrounding discourse)	Morphological future, periphrastic future, present, present subjunctive, periphrastic subjunctive

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

RESULTS

The results from the cross-tabulations reveal the range of forms used by NNSs and NSs to express each potentially variable structure. The frequency with which each form of each dependent variable was used is reported and comparisons are made between the NNSs and the NSs. The section concludes with analyses of frequency of use of each form within the categories of the nonlinguistic independent variables: gender for both groups and time spent abroad and years of language study for the NNSs.

Subject Expression

The analysis of subject positions for all finite verbs revealed categorical use of a null subject in contexts of subject-headed relative clauses (n = 335 for the NNSs and n = 598 for the NSs), atmospheric statements (n = 3 for the NNSs and n = 1 for the NSs), with *hay* "there is" or "there are" (n = 236 for the NNSs and n = 321 for the NSs), and the impersonal or passive *se* (n = 91 for the NNSs and n = 174 for the NSs). Because neither participant group exhibited variation in these syntactic structures, these tokens were eliminated from further analyses.

Table 2 shows the cross-tabulations for the remaining syntactic structures and subject expression use. Null subjects were the most frequent forms for both groups, followed by lexical noun phrases, and then personal pronouns. Demonstrative pronouns, indefinite pronouns, and interrogative pronouns each accounted for less than 5% of use. Categories in which differences of more than 1% are seen include null subject, demonstrative pronoun, and indefinite pronoun. Only the difference in null subject use is greater than 2%.

Table 2. Distribution of subject expression in finite-verb contexts

	NNSs		NSs	
Subject expression	*N*	%	*N*	%
Null subject	3,792	70.2	5,150	67.6
Lexical noun phrase	770	14.3	1,042	13.7
Personal pronoun	589	10.9	793	10.4
Demonstrative pronoun	166	3.1	366	4.8
Indefinite pronoun	61	1.1	229	3.0
Interrogative pronoun	22	0.4	42	0.6
Total	5,400	100	7,622	100

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

Copula Use

In copula-plus-adjective contexts, both groups used *ser* "to be" most often, followed by *estar* "to be." *Parecer* "to seem" and *sentir* "to feel" were less frequent. The results in Table 3 demonstrate that the largest difference in frequency is the 2% difference with *sentir* "to feel." The difference in the use of *ser* "to be" was more than 1%, whereas NNSs and NSs used *parecer* "to seem" and *estar* "to be" at more similar rates, with differences of less than 1%.

Mood Use

In finite-verb contexts, both groups used nonsubjunctive (indicative and nonfinite) forms more often than subjunctive forms.[6] Table 4 also illustrates that the NSs used the subjunctive more than the NNSs, with a difference of 4.7%, and the NNSs used nonsubjunctive forms 4.7% more often than the NSs. As in Gudmestad (2008), both groups produced a range of subjunctive (past, past periphrastic, periphrastic future, pluperfect, present, present perfect) and nonsubjunctive forms (conditional, conditional perfect, conditional periphrastic, imperative, imperfect periphrastic, imperfect progressive, morphological future, morphological future perfect, periphrastic future, pluperfect, present indicative, present perfect, present progressive, preterit, and preterit periphrastic).

Past-Time Reference

In contexts of past-time reference, the present perfect, the preterit, and the imperfect represented 95.1% of the forms used by NNSs and

Table 3. Distribution of copula use in copula-plus-adjective contexts

	NNSs		NSs	
Copula	*N*	%	*N*	%
Estar "to be"	105	17.2	134	17.6
Ser "to be"	479	78.4	607	79.7
Parecer "to seem"	10	1.6	15	2.0
Sentir "to feel"	17	2.8	6	0.8
Total	611	100	762	100

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

Table 4. Distribution of mood use in finite-verb contexts

	NNSs		NSs	
Mood	*N*	%	*N*	%
Subjunctive	124	2.0	586	6.7
Nonsubjunctive	5,941	98.0	8,130	93.3
Total	6,065	100	8,716	100

93.9% of the forms used by NSs. Both groups used the preterit most often, followed by the imperfect, and then the present perfect. The NNSs and the NSs also used the imperfect subjunctive, pluperfect, present, and conditional in contexts of past-time reference, but these forms were used less than 3% of the time by both groups. These findings, summarized in Table 5, also illustrate that the largest differences in use were observed with the present perfect, which the NSs used 4.1% more often than the NNSs, and with the preterit, which the NNSs produced 5.6% more frequently than the NSs. NSs used the past subjunctive and the pluperfect more than the NNSs, with a difference of more than 1%. The differences in rate of use of the imperfect, present, and conditional were less than 1%.

Future-Time Reference

Similar to past-time reference, three forms constituted a majority of participants' use in contexts of future-time reference. The periphrastic

Table 5. Distribution of verb forms used in contexts of past-time reference

	NNSs		NSs	
Verb form	*N*	%	*N*	%
Present perfect	128	17.3	295	21.4
Preterit	323	43.6	523	38.0
Imperfect	254	34.3	475	34.5
Past subjunctive	4	0.5	24	1.7
Pluperfect	9	1.2	39	2.8
Present	18	2.4	20	1.5
Conditional	5	0.7	1	0.1
Total	741	100	1,377	100

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

future, present, and morphological future comprised 97.2% of NNS use and 88.4% of NS use. Both groups used the periphrastic future most frequently. The NNSs used the present more often than the morphological future, whereas the NSs used the morphological future more frequently than the present. The conditional, present subjunctive, and periphrastic subjunctive were used less than 3% of the time, with the exception of the present subjunctive, which NSs produced 9% of the time, and the periphrastic subjunctive, which the NNSs did not use. Additionally, Table 6 shows that the largest difference between groups was with the periphrastic future; the NNSs used this form 16.9% more often than the NSs. The NSs used the morphological future, the present subjunctive, and the periphrastic future more frequently than the NNSs (differences of 7.7%, 6.7%, and 2.3%, respectively). The differences in the rate of use of the present and the conditional were less than 1%.

Nonlinguistic Variables

Cross-tabulations were carried out for the individual variables, and the rate of use of each category of each dependent variable (e.g., subject expression) across the categories of each independent variable (e.g., female gender) was compared with the average rates of use. The discussion of key similarities and differences among the categories of nonlinguistic variables is guided by the definition of *similar* as indicating a difference of less than 2% from the overall average rate of use of a given form, whereas *different* indicates a difference of 2% or more. These qualitative comparisons serve to begin the exploration of the role that these individual characteristics may play in

Table 6. Distribution of verb forms used in contexts of future-time reference

	NNSs		NSs	
Verb form	*N*	%	*N*	%
Morphological future	19	8.8	57	16.5
Periphrastic future	164	75.9	204	59.0
Present	27	12.5	45	13.0
Conditional	1	0.5	1	0.3
Present subjunctive	5	2.3	31	9.0
Periphrastic subjunctive	0	0.0	8	2.3
Total	216	100	346	100

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

the use of the five potentially variable structures. The results of each of the cross-tabulations are presented in percentages in Table 7.

Gender. Beginning with the NNSs, differences of 2% or more between females or males and the average for each category of each dependent variable were sometimes observed in copula use, past-time reference, and future-time reference but not in any category of subject expression or mood use. In copula-plus-adjective contexts, the NNSs used *estar* "to be" 17.2% of the time. When analyzed separately, the NNS males used *estar* "to be" (58/297 tokens) more frequently than the NNS females (47/314 tokens). In contexts of past-time reference, NNSs used the present perfect 17.3% of the time. Males used this form (87/450 tokens) more than females (41/291 tokens). However, the average for the preterit was 43.6%, and the females produced this form (140/323 tokens) more often than the males (183/450 tokens). Finally, in contexts of future-time reference, NNSs' average rate of use of the periphrastic future was 75.9%. The males produced this structure (113/145 tokens) more often than the females (51/71 tokens). The females differed more than two percentage points (11/27 tokens) from the average rate of the use of the present (12.5%), but the males did not.

For the NSs, differences were seen in some categories of subject expression, past-time reference, and future-time reference. For subject expression, the average rate of null-subject use was 67.6%. Males used null subjects (2,734/3,847 tokens) more frequently than females (2,416/3,775 tokens). In contrast, the average rate of personal-pronoun use was 10.4%, and females (527/3,775 tokens) used more personal pronouns than males (266/3,847 tokens). In contexts of past-time reference, the average rate of present perfect use was 21.4%. Males used this form (149/606 tokens) more often than females (146/317 tokens). Furthermore, the females used the preterit (317/771 tokens) more than the males (206/606 tokens), and the average was 38%. Finally, in contexts of future-time reference, the males used the morphological future (37/185 tokens) more than the females (20/161 tokens), and the average was 16.5%. The average rate of use for the periphrastic future was 59.0%; the females (103/161 tokens) used this structure more frequently than the males (101/185 tokens).

Time Abroad. To carry out the cross-tabulation for time spent in a Spanish-speaking country, the NNSs were divided into two groups: those who had spent less than 1 year abroad (N = 8) and those who had spent 1 year or more abroad (N = 8). This decision was made because the 1-year mark generally separates learners who participate in academic programs abroad from those who spend more time for purposes such as extended study or work opportunities. Differences of

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

Table 7. Percent of occurrence of each category of each dependent variable across categories of independent nonlinguistic variables

	Gender				Time abroad		Years of study	
Dependent variable	Female NNSs	Male NNSs	Female NSs	Male NSs	<1 year	≥1 year	<9 years	≥9 years
Subject expression								
Null	70.6	70.0	64.0	71.1	71.7	69.1	70.3	70.2
Lexical noun phrase	14.8	13.9	12.7	14.6	14.2	14.3	14.3	14.2
Personal pronoun	11.1	10.8	14.0	6.9	9.8	11.8	11.7	10.1
Demonstrative pronoun	2.3	3.6	5.1	4.5	3.2	3.0	2.1	4.1
Indefinite pronoun	0.9	1.3	3.4	2.6	0.8	1.4	1.3	1.0
Interrogative pronoun	0.4	0.4	0.7	0.4	0.3	0.5	0.4	0.4
Copula use								
Estar "to be"	15.0	19.5	18.0	17.1	14.4	19.9	17.9	16.4
Ser "to be"	80.9	75.8	81.2	78.2	80.0	76.8	78.0	78.9
Parecer "to be"	0.6	2.7	0.5	3.4	2.0	1.3	2.9	2.7
Sentir "to be"	3.5	2.0	0.3	1.3	3.6	2.0	1.3	2.0
Mood use								
Subjunctive	1.5	2.4	7.3	6.1	1.6	2.4	2.3	1.8
Nonsubjunctive	98.5	97.6	92.7	93.9	98.4	97.6	97.2	98.2
Past-time reference	14.1	19.3	18.9	24.6	14.6	19.3	15.0	20.1
Present perfect	48.1	40.7	41.1	34.0	40.8	45.6	41.0	46.8
Preterit	33.7	34.7	35.1	33.7	38.0	31.5	39.6	27.7
Imperfect	0.0	0.9	1.7	1.8	0.3	0.7	0.2	0.9
Past subjunctive	0.7	1.6	1.9	4.0	1.6	0.9	1.0	1.5
Pluperfect	2.7	2.2	1.2	1.8	4.4	0.9	2.4	2.4
Present conditional	0.7	0.7	0.0	0.2	0.3	0.9	0.7	0.6
Future-time reference	8.5	9.0	12.4	20.0	11.2	6.8	3.6	14.3
Morphological future	71.8	77.9	64.0	54.6	69.4	81.4	79.3	72.4
Periphrastic future	15.5	11.0	11.2	14.6	17.3	8.5	13.5	11.4
Present conditional	0.0	0.7	0.0	0.5	0.0	0.8	0.9	0.0
Present subjunctive	4.2	1.4	9.3	8.6	2.0	2.5	2.7	1.9
Periphrastic subjunctive	0.0	0.0	3.1	1.6	0.0	0.0	0.0	0.0

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

2% or more were observed with some categories of copula use, past-time reference, and future-time reference. For the copula *estar* "to be" (which had an overall average rate of use among NNSs of 17.2%), NNSs who had studied abroad for at least 1 year used *estar* "to be" (61/306 tokens) more often than those who had spent less than 1 year abroad (44/305 tokens). In contexts of past-time reference, participants who had spent more than 1 year abroad produced the present perfect (82/425 tokens) and the preterit (194/425 tokens) more frequently than those who had studied abroad for less than 1 year (present perfect: 46/316 tokens; preterit: 129/316 tokens). The overall average rates of use of these forms among NNSs were 17.3% and 43.6%, respectively. In contrast, for the imperfect (which had an average rate of use of 34.5% among the NNSs) participants who had spent less time abroad used this form (120/316 tokens) more often than those who had been abroad for 1 year or more (134/425 tokens). Finally, differences between the morphological future (8.8% overall average rate of use among NNSs), periphrastic future (75.9% average), and present (12.5% average) were also observed. The speakers who had been abroad for 1 year or more used more periphrastic future (96/118 tokens) than those with less time spent abroad (68/98 tokens). However, the participants with less than 1 year abroad used more morphological future (11/98 tokens) and present (17/98 tokens) than those with 1 year or more in a Spanish-speaking country (morphological future: 8/118 tokens; present: 10/118 tokens).

Years of Study. The NNSs were divided into two groups for the analysis of years of formal study of Spanish: those who had studied for 9 years or more (N = 8) and those who had studied for less than 9 years (N = 8). Because the mean length of study was just over 8 years overall, the 9-year cutoff point was selected to maintain an even number of participants in each group. Differences of 2% or more were observed with some categories of past-time reference and future-time reference. In contexts of past-time reference, NNSs who had studied Spanish longer used more present perfect (66/329 tokens) and preterit (154/329 tokens) than those who had studied for less than 9 years (present perfect: 62/412 tokens; preterit: 169/412 tokens). The overall average rates of use of these forms among NNSs were 17.3% and 43.6%, respectively. In contrast, the imperfect, with an average rate of use of 34.5%, was used more often by those who had studied Spanish for less than 9 years (163/412 tokens) than by those who had studied for more time (91/329 tokens). Finally, in contexts of future-time reference, the NNSs with more years of language study used more morphological future (15/105 tokens) than those with fewer years of study (4/111 tokens); the overall average was 8.8%. The periphrastic future, which had an average rate of use of 75.9%,

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

was used more frequently by the NNSs who had studied Spanish for less than 9 years than by those who had studied for 9 years or more (76/105 tokens).

The results reveal differences between NNSs and NSs in the frequency of forms in the five potentially variable structures. First, although all groups produced overt subjects (noun phrases or pronouns) about 30% of the time, the NSs used slightly more overt subjects than the NNSs. Second, both groups used *ser* "to be" and *estar* "to be" at similar rates but differed in their use of *parecer* "to seem" and *sentir* "to feel." Third, NNSs produced the subjunctive less often than the NSs. Fourth, the imperfect, preterit, and present perfect were most likely to be used in contexts of past-time reference. Whereas frequency of imperfect was equal for NSs and NNSs, NSs used more present perfect and less preterit than NNSs. Fifth, in contexts of future-time reference, present tense was produced with equal frequency for both groups, whereas NSs used more morphological future forms and fewer periphrastic future forms than NNSs. Finally, regarding the nonlinguistic variables, differences were observed for gender, time spent abroad, and years of formal study. NNS males used more *estar* "to be," present perfect, and periphrastic future, whereas NNS females produced more preterit and present (in contexts of future-time reference). NS males produced more null subjects, present perfect, and morphological future, whereas the NS females used more personal pronouns, preterit, and periphrastic future. Moreover, the NNSs who had studied abroad for 1 year or more used more *estar* "to be," present perfect, preterit, and periphrastic future than those with less study-abroad experience. The NNS participants who had studied in a Spanish-speaking country for less than 1 year used more imperfect, morphological future, and present (in contexts of future-time reference) than the other NNS group. Finally, the NNSs who had studied Spanish for at least 9 years used more present perfect, preterit, and morphological future than those with fewer years of formal study, who produced more imperfect and periphrastic future than the other group.

DISCUSSION

Range of Forms Used

The analysis of five potentially variable structures in Spanish, in which the definition of the context itself is the key expansion of previous research, has revealed that a broader range of forms than has traditionally been analyzed occurs in these contexts, except in the case of mood use. This broader range of forms, in addition to addressing the challenge of identifying contexts in which the same meaning is conveyed, demonstrates that the predictive models typically employed in variationist

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

studies are inappropriate for these structures, at least until sufficient preliminary analyses have been conducted. Moreover, excluding some learner forms can mask components of interlanguage and provide an incomplete and misleading characterization of the developing system. Furthermore, largely the same range of forms was found for the NSs, which challenges the approach taken in much of the existing research on these five structures. Previously held assumptions that variation in these five structures is limited to only two categories of the dependent variable should be reconsidered. The current study does not claim that variationist analyses are impossible but rather that the research presented here is a necessary first step to fully understand language use by NSs or NNSs. A likely second step may involve an analysis of the predictive factors that govern a subset of the data for which a binary dependent variable exists. This approach, of course, will require researchers to either address whether the conditions for considering these structures to be linguistic variables have been met or to accept that a single function can be the object of analysis (e.g., Schwenter & Torres Cacoullos, 2008). For example, the current analysis showed that for past-time forms, both NSs and NNSs used the imperfect with similar frequency, whereas the use of preterit and present perfect differed across groups. Once these factors have been identified and when it has been shown that the more stable form in terms of frequency appears in the same types of linguistic contexts for both groups, it will be possible to examine only the contexts in which the two seemingly more variable forms occur and through which the same meaning is conveyed (e.g., past and completed event) or the same function is met. The present study has also shown that the traditional sociolinguistic predictive analysis may not be the immediate goal for some potentially variable structures, such as subject expression, because language use is complex enough that many issues must be resolved before such an undertaking would be appropriate. The findings support the view of interlanguage as a complex system in which several forms are interrelated and the contexts of occurrence of these forms must be more carefully identified and defined. This type of approach should probably be applied to sociolinguistic research as well, which could take on more rigorous methodology in defining the envelope of variation prior to eliminating particular subsets of the data where variation is believed not to occur.

NS-NNS Comparison

Even though NNSs and NSs produced a similar range of forms to express the five potentially variable structures investigated, there are notable differences between these groups. First, the frequency with which

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

NNSs used many categories of each dependent variable was clearly different from the NS group. This observation demonstrates that obligatory contexts are not needed to identify differences between groups. Moreover, such differences lead to hypotheses that can be tested in future research. For example, the fact that NNSs used the subjunctive noticeably less than the NSs may indicate that the NSs produced more appropriate contexts for the subjunctive or may have employed certain linguistic triggers for the subjunctive with greater frequency—a hypothesis that can be tested in future studies. Additionally, in looking at future-time expression, both groups used the present indicative with similar frequency; the main difference in use was found between the periphrastic and morphological futures. This type of variation is precisely what has been shown to occur across geographic regions and to be linked to variation in linguistic features of the context (e.g., Blas Arroyo, 2008; Orozco, 2005). Thus, it would be reasonable to analyze this difference between the groups using the tools of variation analysis normally employed in sociolinguistic research to determine whether both groups use these forms in response to the same linguistic features. These two examples show that two contributions of the current study are the identification of differences between NSs and NNSs as well as testable hypotheses about the source(s) of these differences.

Nonlinguistic Variables

The results may suggest that the NNSs do not exhibit nativelike gender variation. Differences were observed between NS females and males with null subjects, personal pronouns, and morphological future but not for the NNSs; the opposite trend of use was observed between genders of each speaker group with the periphrastic future, and differences were observed with the NNSs that were not seen with the NSs (*estar* "to be" and the use of present in contexts of future-time reference). However, similar gender patterns were observed across speaker groups for the present perfect and preterit. These findings remain tentative due to the small participant groups and because further details about the types of contexts in which these forms occur may reveal that some differences are attributable to other factors (e.g., variation in types of contexts produced). Nevertheless, the findings suggest that gender is a social factor worthy of further attention.

The current study also demonstrates that time abroad and length of formal study of Spanish contribute to the use of L2 forms in potentially variable structures. The NNSs who had spent 1 year or more abroad used more *estar* "to be," present perfect, preterit, and periphrastic future and produced fewer imperfect, morphological future, and present

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

(in future-time contexts) than NNSs with less time spent abroad. The NNSs who had studied Spanish for 9 years or more also produced more present perfect, preterit, and morphological future forms as well as fewer imperfect and periphrastic future forms. The results indicate that even very advanced learners demonstrate differences based on the degree to which they have been exposed to input, both in formal and naturalistic contexts. This finding underscores the importance of examining very advanced learners to study the ways in which they approximate NS use and highlights the value of the current analysis, which identifies the relationship between individual learner characteristics and L2 use.

CONCLUSIONS AND FUTURE DIRECTIONS

The current study showed that relatively advanced learners of Spanish and NSs produce the same range of forms in contexts of subject expression, copula contrast, mood distinction, past-time reference, and future-time reference but that these forms occur at different rates. The data also demonstrated that individual characteristics are related to language use. Social variation linked to gender was identified as an area of interest for future research, and exposure to Spanish, as indicated by time spent in a Spanish-speaking country and years of formal study, contributed to the NNSs' use of certain forms of the potentially variable structures.

This large-scale study of potentially variable structures in Spanish lays the foundation for future work. First, in-depth analyses of specific structures are needed. This study has examined a range of forms and the frequency of their use, but the specific linguistic contexts in which these forms are used must be examined to learn if these potential variants meet the conditions of the traditional linguistic variable. Second, this study compares one L2 group to NSs but does not make any observations about development. Research that examines cross-sectional or longitudinal data is needed to identify stages of development of these structures both in terms of the range of forms used and the linguistic and extralinguistic factors that characterize this use (for developmental stages of mood use, see Gudmestad, 2008, and for copula contrast see Geeslin, 2005). Third, previous research has not empirically defined the envelope of variation (e.g., Cameron, 1994) for these potentially variable structures; therefore, studies that examine this issue for NSs and learners are needed. Finally, further exploration of the link between the study-abroad context and the SLA of variable structures is encouraged, to identify the stage at which learners begin to show variation and whether this shows characteristics of the variety with which the learners have contact or reflects their own social characteristics.

(*Received 26 June 2009*)

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

NOTES

1. Sankoff (1988) addressed grammatical functions for which form distinctions can be neutralized in discourse, and Schwenter and Torres Cacoullos (2008) noted that forms that serve the same function cannot be defined narrowly because they cover a range of meanings along the grammaticalized path. The path itself and these meanings constitute the so-called envelope of variation (e.g., Cameron, 1994).

2. Second-language studies do not generally address the debate of considering nonphonological structures to be linguistic variables and there are many examples of the extension of this analysis (Gudmestad, 2006; Young, 1991).

3. The term *distinction* denotes the subjunctive-indicative alternation, whereas the term *use* is restricted to language production and the terms *choice* and *selection* refer to tasks with response options.

4. All participants had the same opportunities to produce the same contexts for each variable, but this does not mean that all participants produced the same contexts. Each participant produced tokens according to the meaning they chose to express in response to each question.

5. In sociolinguistics and SLA research, chi-square tests are expected with cross-tabulations. Nevertheless, in the current study, the analysis is not focused on such tests, given the common assumptions in the field of statistics with regard to token independence. It is expected, however, that some readers will prefer to see such tests so as to relate this work to its predecessors. Thus, it should be noted that the chi-square tests for participant group and each dependent variable yielded significant differences between groups. For the independent nonlinguistic variables, significant differences between genders were found for mood use for NNSs and NSs and for subject expression, copula use, and past-time reference for NSs. There were significant relationships between time spent abroad and past-time reference and between years of study and subject expression and past-time reference.

6. The total number of tokens per group in Table 4 (subjunctive use) differs from Table 2 (subject expression) because the analysis of mood includes contexts of categorical subject-expression use.

REFERENCES

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

Andersen, R. W. (1991). Developmental sequences: The emergence of aspect marking in second language acquisition. In T. Huebner & C. A. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theories* (pp. 305–324). Amsterdam: Benjamins.

Andersen, R. W., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition*, *16*, 133–156.

Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning, and use*. Oxford: Blackwell.

Bardovi-Harlig, K. (2004). Monopolizing the future: How the go-future breaks into will's territory and what it tells us about SLA. *EUROSLA Yearbook*, *4*, 177–201.

Bayerová, M. (1994). Alternancia indicativo × subjuntivo en oraciones independientes [Subjunctive and indicative alternation in independent sentences]. *Etudes Romanes de Brno*, *24*, 61–71.

Bayley, R. (1996). Competing constraints on variation in the speech of adult Chinese learners of English. In R. Bayley & D. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 97–130). Amsterdam: Benjamins.

Bayley, R., & Langman, J. (2004). Variation in the group and the individual: Evidence from second language acquisition. *IRAL*, *42*, 303–318.

Bayley, R., & Pease-Álvarez, L. (1997). Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change*, *9*, 349–371.

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

Blas Arroyo, J. L. (2008). The variable expression of future tense in Peninsular Spanish: The present (and future) of inflectional forms in Spanish spoken in a bilingual region. *Language Variation and Change*, *20*, 85–126.
Blas Arroyo, J. L., & Porcar Miralles, M. (1997). Aproximación sociolingüística al fenómeno de la neutralización modal en las comunidades de habla castellonenses (análisis de algunos contornos sintácticos) [Sociolinguistic approximation of the phenomenon of modal neutralization in Spanish-speaking communities (analysis of some syntactic contexts)]. *Sintagma*, *9*, 27–45.
Brugger, G. (2001). Temporal modification, the 24-hour rule and the location of reference time. In J. Gutiérrez-Rexach & L. Silva-Villar (Eds.), *Current issues in Spanish syntax and semantics* (pp. 243–270). Berlin: Mouton de Gruyter.
Butt, J., & Benjamin, C. (2004). *A new reference grammar of modern Spanish* (4th ed.). New York: McGraw Hill.
Cadierno, T. (2000). The acquisition of Spanish grammatical aspect by Danish advanced language learners. *Spanish Applied Linguistics*, *4*, 1–53.
Cameron, R. (1994). Switch reference, verb class and priming in a variable syntax. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, & E. Zeinfeld (Eds.), *Papers from the 30th regional meeting of the Chicago Linguistics Society: Vol. 2. The parasession on variation in linguistic theory* (pp. 27–45). Chicago: Chicago Linguistics Society.
Cameron, R., & Flores-Ferrán, N. (2004). Perseverance of subject expression across regional dialects. *Spanish in Context*, *1*, 41–65.
Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*, 1–47.
Cheshire, J. (1987). Syntactic variation, the linguistic variable, and sociolinguistic theory. *Linguistics*, *25*, 257–282.
Clements, J. C. (2005). 'Ser' and 'estar' in the predicate adjective construction. In J. C. Clements & J. Yoon (Eds.), *Functional approaches to Spanish syntax: Lexical semantics, discourse, and transitivity* (pp. 161–202). London: Palgrave-Macmillan.
Collentine, J. (2003). The development of the subjunctive and complex-syntactic abilities among foreign language learners of Spanish. In B. A. Lafford & M. R. Salaberry (Eds.), *Spanish second language acquisition: State of the science* (pp. 74–97). Washington, DC: Georgetown University Press.
Cortés-Torres, M. (2004). ¿Ser o estar? La variación lingüística y social de estar más adjetivo en el español de Cuernavaca, Mexico [Ser or estar? Linguistic and social variation of estar plus adjective in the Spanish of Cuernavaca, Mexico]. *Hispania*, *87*, 788–795.
Díaz-Campos, M., & Geeslin, K. L. (in press). Copula use in the Spanish of Venezuela: Social and linguistic sources of variation. *Spanish in Context*.
Ellis, R. (1999). Item versus system learning: Explaining free variation. *Applied Linguistics*, *20*, 460–480.
Fernández Leborans, M. J. (1999). La predicación: Las oraciones copulativas [Predication: Copular sentences]. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (pp. 2354–2460). Madrid: Espasa.
Geeslin, K. L. (2003). A comparison of copula choice in advanced and native Spanish. *Language Learning*, *53*, 703–764.
Geeslin, K. L. (2005). *Crossing disciplinary boundaries to improve the analysis of second language data*. Munich: Lincom Europa.
Geeslin, K. L., & Gudmestad, A. (2008). Comparing interview and written elicitation tasks in native and non-native data: Do speakers do what we think they do? In J. Bruhn de Garavito & E. Valenzuela (Eds.), *Selected proceedings of the 10th Hispanic Linguistics Symposium* (pp. 64–77). Somerville, MA: Cascadilla Press.
Geeslin, K. L., & Guijarro-Fuentes, P. (2006). The second language acquisition of variable structures in Spanish by Portuguese speakers. *Language Learning*, *56*, 53–107.
Geeslin, K. L., & Guijarro-Fuentes, P. (2008). Variation in contemporary Spanish: Linguistic predictors of *estar* in four cases of language contact. *Bilingualism: Language and Cognition*, *11*, 365–380.
Gudmestad, A. (2006). L2 variation and the Spanish subjunctive: Linguistic features predicting use. In C. A. Klee & T. Face (Eds.), *Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 170–184). Somerville, MA: Cascadilla Press.

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

Gudmestad, A. (2008). Acquiring a variable structure: An interlanguage analysis of second-language mood use in Spanish (Doctoral dissertation, Indiana University, 2008). *Dissertation Abstracts International*, *69*, 8.

Gutiérrez, M. (1995). On the future of the future tense in the Spanish of the Southwest. In C. Silva-Corvalán (Ed.), *Spanish in four continents: Studies in language contact and bilingualism* (pp. 214–226). Washington, DC: Georgetown University Press.

Gutiérrez, M. (2003). Simplification and innovation in US Spanish. *Multilingua*, *22*, 169–184.

Haverkate, H. (2002). *The syntax, semantics and pragmatics of Spanish mood*. Amsterdam: Benjamins.

Isabelli, C. A. (2004). The acquisition of the null subject parameter properties in SLA: Some effects of positive evidence in a naturalistic learning context. *Hispania*, *87*, 150–162.

Klee, C. A., & Ocampo, A. M. (1995). The expression of past tense referent in Spanish narratives of Spanish-Quechua bilingual speakers. In C. Silva-Corvalán (Ed.), *Spanish in four continents: Studies in language contact and bilingualism* (pp. 52–70). Washington, DC: Georgetown University Press.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, W. (1978). *Where does the linguistic variable stop? A response to Beatriz Lavandera*. (Working Papers In Sociolinguistics 44). Austin, TX: Southwest Educational Development Laboratory.

Lafond, L., Hayes, R., & Bhat, R. (2000). Constraint demotion and null-subjects in Spanish L2 acquisition. In J. Camps & C. Wiltshire (Eds.), *Romance syntax, semantics and L2 acquisition: Selected papers from the 30th Linguistic Symposium on Romance Languages* (pp. 121–135). Amsterdam: Benjamins.

Lee, J. F. (2002). The incidental acquisition of Spanish: Future tense morphology through reading in a second language. *Studies in Second Language Acquisition*, *24*, 55–80.

Liskin-Gasparro, J. (2000). The use of tense-aspect morphology in Spanish oral narratives: Exploring the perceptions of advanced learners. *Hispania*, *83*, 830–844.

López Ortega, N. (2003). The development of discourse competence in study abroad learners: A study of subject expression in Spanish as a second language. (Doctoral dissertation, Cornell University, 2003). *Dissertation Abstracts International*, *63*, 9.

Lozano, C. (2002). Knowledge of expletive and pronominal subjects by learners of Spanish. *ITL Review of Applied Linguistics*, *135*, 37–60.

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

Mackenzie, I. (1995). The supposed imperfectivity of the Latin American present perfect. *Hispanic Linguistics*, *6*, 29–60.

Mesa Alonso, M., Domínguez Herrera, M., Padrón Sánchez, E., & Morales Aguilera, N. (1993). Ser y estar: Consideraciones sobre su uso en español [Ser and estar: Considerations of use in Spanish]. *Islas*, *104*, 150–156.

Murillo Medrano, J. (1999). Subjuntivo e indicativo en las oraciones circunstanciales [Subjunctive and indicative in circumstantial sentences]. *Kanina*, *23*, 143–156.

Nadasdi, T., Mougeon, R., & Rehner, K. (2003). The use of the "future tense" in the spoken French of French immersion students. *Journal of French Language Studies*, *13*, 195–219.

Orozco, R. (2005). Distribution of future time forms in Northern Colombia Spanish. In D. Eddington (Ed.), *Selected proceedings of the 7th Hispanic Linguistics Symposium* (pp. 56–65). Somerville, MA: Cascadilla Press.

Otheguy, R., & Zentella, A. (2007). Apuntes preliminares sobre el contacto lingüístico y dialectal en el uso pronominal del español en Nueva York [Preliminary notes on linguistic and dialect contact on pronoun use in Spanish in New York]. In K. Potowski & R. Cameron (Eds.), *Spanish in contact: Policy, social and linguistic inquiries* (pp. 275–298). Amsterdam: Benjamins.

Preston, D. (2000). Three kinds of sociolinguistics and SLA: A psycholinguistic perspective. In B. Swerzbin, F. Morris, M. Anderson, C. Klee, & E. Tarone (Eds.), *Social and cognitive factors in SLA* (pp. 3–30). Somerville, MA: Cascadilla Press.

Regan, V. (1996). Variation in French interlanguage: A longitudinal study of sociolinguistic competence. In R. Bayley & D. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 177–201). Amsterdam: Benjamins.

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press

Rehner, K. (2002). The development of aspects of linguistic and discourse competence by advanced second language learners of French. (Doctoral dissertation, University of Toronto, 2002). *Dissertation Abstracts International*, *63*, 12.
Rehner, K., Mougeon, R., & Nadasdi, T. (2003). The learning of sociolinguistic variation by advanced FSL learners: The case of *nous* versus *on* in immersion French. *Studies in Second Language Acquisition*, *25*, 127–156.
Romaine, S. (1984). On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory. *Folia Lingüística*, *18*, 409–437.
Rossomondo, A. E. (2007). The role of lexical temporal indicators and text interaction format in the incidental acquisition of the Spanish future tense. *Studies in Second Language Acquisition*, *29*, 39–66.
Ryan, J., & Lafford, B. A. (1992). The acquisition of lexical meaning in a study abroad environment: *Ser* + *estar* and the Granada experience. *Hispania*, *75*, 714–722.
Salaberry, M. R. (2008). *Marking past tense in second language morphology: A theoretical model*. New York: Continuum.
Salaberry, M. R., & Shirai, Y. (Eds.). (2002). *The L2 acquisition of tense-aspect morphology*. Amsterdam: Benjamins.
Sankoff, D. (1988). Sociolinguistics and syntactic variation. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey* (Vol. 4, pp. 140–161). New York: Cambridge University Press.
Sankoff, G., & Vincent, D. (1977). L'emploi productif de 'ne' dans le français parlé de Montréal [The productive use of 'ne' in the spoken French of Montreal]. *Le Français Moderne*, *45*, 243–256.
Schwenter, A., & Torres Cacoullos, R. (2008). Defaults and interdeterminacy in temporal grammaticalization: The 'perfect' road to perfective. *Language Variation and Change*, *20*, 1–39.
Silva-Corvalán, C. (1982). Subject expression and placement in Mexican-American Spanish. In J. Amastae & E. Olivares (Eds.), *Spanish in the United States: Sociolinguistic aspects* (pp. 93–120). New York: Cambridge University Press.
Silva-Corvalán, C. (1994a). The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change*, *6*, 255–272.
Silva-Corvalán, C. (1994b). *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
Silva-Corvalán, C., & Terrell, T. (1989). Notas sobre la expresión de futuridad en el español del Caribe [Notes on the expression of futurity in Caribbean Spanish]. *Hispanic Linguistics*, *2*, 191–208.
Tarone, E. (2007). Sociolinguistic approaches to second language acquisition research: 1997–2007. *Modern Language Journal*, *91*, 837–848.
Vañó-Cerdá, A. (1982). *Ser y estar* + *adjetivos* [Ser and estar + adjectives]. Tübingen: Gunter Narr Verlag.
VanPatten, B. (1987). The acquisition of ser and *estar*: Accounting for developmental patterns. In B. VanPatten, T. Dvorak, & J. F. Lee (Eds.), *Foreign language learning: A research perspective* (pp. 61–75). Rowles, MA: Newbury House.
Weiner, J., & Labov, W. (1983). Constraints on agentless passive. *Journal of Linguistics*, *19*, 29–58.
Winford, D. (1996). The problem of syntactic variation. In J. Arnold, R. Blake, B. Davidson, S. Schwenter, & J. Solomon (Eds.), *Sociolinguistic variation: Data, theory, and analysis—Selected papers from NWAV 23 at Stanford* (pp. 177–192). Stanford, CA: Center for the Study of Language and Information.
Wolfram, W. (1985). Variability in tense marking: A case for the obvious. *Language Learning*, *35*, 229–253.
Young, R. (1991). *Variation in interlanguage morphology*. Bern: Peter Lang.

https://doi.org/10.1017/S0272263110000033 Published online by Cambridge University Press