Research Article ⬡

SOCIOLINGUISTIC COMPETENCE AND INTERPRETING VARIABLE STRUCTURES IN A SECOND LANGUAGE

A STUDY OF THE COPULA CONTRAST IN NATIVE AND SECOND-LANGUAGE SPANISH

*Matthew Kanwit** ⬥

University of Pittsburgh

Kimberly L. Geeslin

Indiana University

Abstract
Learners must develop the ability to vary language according to linguistic and situational factors to produce context-appropriate utterances. Likewise, interpreting the additional meaning conveyed through language variation is essential for successful communication. Nevertheless, research on the interpretation of the variable copulas in Spanish is scarce and we do not know how individual lexical items and patterns of co-occurrence of adjectives with particular copulas influence interpretation. Addressing this void, we compare interpretation of the copulas by native speakers and highly advanced, advanced, and intermediate learners. Participants completed an interpretation task containing the copulas paired with one of nine adjectives, categorized as typically co-occurring with *ser*, *estar*, or both copulas. The current study contributes to the body of work on communicative competence and advanced L2 proficiency by exploring the development of interpretative abilities of English-speaking learners of Spanish and the extent to which interpretation differs across adjective classes and individual adjectives.

Language learners must develop the ability to vary language according to linguistic and situational factors to produce context-appropriate utterances. Likewise, interpreting the additional meaning conveyed through language variation to interact effectively across

multiple interactional contexts is essential. The study of communicative competence (e.g., Canale & Swain, 1980; Sun, 2014; van Compernolle & Williams, 2012) is long-standing and was initially explored to understand what language teachers must do beyond assisting learners in the development of grammatical competence to foster the acquisition of language abilities through which learners could participate fully in the target language. Over the past decades, this field of research has also led to an increasingly large and sophisticated body of work demonstrating how factors specific to the linguistic components of a language, the characteristics of individual speakers, and the context of interaction all come to influence the production and interpretation of the variable elements in language in context. Much of this work has been conducted within usage-based, functional, and variationist approaches to language acquisition and use. Within these frameworks, we see that patterns of language use drive acquisition and that these patterns are influenced by factors ranging from how often a given lexical item occurs in a language to the social identities of the speaker and hearer. In the context of second language (L2) acquisition, we see that even if production of variable structures indicating particular social, situational, or geographic characteristics is optional for individual speakers, targetlike interpretation of these forms is integral to advanced linguistic competence. The study contributes to this body of work through an examination of the copula contrast in Spanish, a variable structure in native and L2 populations, by exploring the development of interpretative abilities of English-speaking learners of Spanish, and the extent to which interpretation is differential across adjective classes and individual adjectives.

THE RESEARCH CONTEXT

The current investigation lies at the intersection of several subfields of the study of language acquisition and use. It is contextualized within usage-based frameworks that recognize language variation and the multiple factors that influence these variable patterns of use. Likewise, the study employs a variationist analysis and is part of a growing body of research on L2 variation. We focus specifically on interpretative abilities, an area of study that remains relatively small. Finally, the test case to which we apply our research questions is the copula contrast in L2 Spanish, and this is among the structures about which we have the most empirical information in the Spanish language, both among native speakers (NSs) and L2 learners. In the review that follows, we summarize relevant research in these domains to contextualize the study and identify the contributions that our research makes for each of these areas of investigation.

USAGE-BASED APPROACHES TO LANGUAGE PRODUCTION AND ACQUISITION

Generally speaking, usage-based approaches are characterized by their attention to patterns attested in the language to which a speaker is exposed. These patterns drive language acquisition and contextualize language structures within their linguistic, social, and interactional settings (e.g., Bybee, 2007, 2010; Goldberg, 2013). Likewise, usage-based approaches view language as a series of connections between lexical, structural, and social factors, rather than as an isolated entity. These approaches attend to the role of the frequency of occurrence of particular lexical items, or collocations thereof, in

processes of language change and language acquisition. Research shows that highly frequent grammatical items resist changes that result from processes such as analogy, whereas less frequent forms are less resistant to such change. Thus, language change is diffused through the lexicon differentially, affecting items to varying extents based on their overall frequency. Moreover, the patterns that are the most lexically diffuse (i.e., that extend to the greatest number of lexical items) are more likely to attract members from other classes in a type of pattern regularization. In this sense, both lexical (i.e., token) frequency and type frequency play a role in acquisition. In contrast to patterns of syntactic change, phonological change is generally led by the more frequent forms, which tend to reduce more often and to greater degrees (Bybee, 2007, 2010). Despite the importance of frequency in these approaches, it is a valuable starting point for the current discussion to note that most usage-based approaches also allow for a level of abstraction. Thus, it is not necessary to argue that frequency, or any other factor, is the sole determiner of change or acquisition, nor that languages operate in the absence of generalization.

Usage-based approaches have been applied profitably to the study of L2 acquisition. This body of work demonstrates the importance of constructs such as lexical frequency, collocational frequency, lexical density, and lexical deployment and their relationship to L2 learning. For example, we know that learners exhibit sensitivity to lexical frequency in a variety of ways. Ellis and Schmidt (1997) showed that learners of an artificial grammar were more accurate with past tense forms for high-frequency verbs, and Wulff, Ellis, Römer, Bardovi-Harlig, and Leblanc (2009) showed that those verb forms that occurred most frequently in the progressive and past tense in their NS corpus were the first to emerge in those same forms in the L2 of adult learners. Ellis and Ferreira-Junior (2009) studied verb/argument constructions among native and L2 speakers of English, showing that those forms that occur most frequently among NSs were even more strongly represented by the nonnatives and were often the first to emerge. The implication is that frequency may lead to prototypes that foster the development within a particular category. Durrant and Schmitt (2009) examined collocation type, where collocation is defined as a relationship that a lexical item has with items that co-occur with greater than random probability (following Hoey, 1991, p. 7), and showed that learners produced highly frequent collocations at an even higher rate than NSs while underusing collocations with lower frequencies of occurrence. Crossley, Subtirelu, and Salsbury (2013) identified word frequency as the strongest predictor in early noun production in an L2 (see also Wolter & Gyllstad, 2013 for similar work on collocations). These results have been further corroborated across several tasks, showing that frequency affects speed of reading, accuracy in acceptance on grammaticality judgment tasks, spoken articulation, the degree to which reading the beginning of the formula primed recognition of the second word (i.e., primed production), and speed of comprehension (e.g., Ellis, Simpson-Vlach, & Maynard, 2008). Recent research on language processing has also shown that the facilitative effects among bilinguals that result from higher frequency are greater in the L2 than the first language (L1), perhaps due to smaller vocabulary size as a result of less exposure to the target (Diependaele, Lemhöfer, & Brysbaert, 2013; see also Yi, 2018, for collocations) and that this effect is not mediated by L2 proficiency (Cop, Keuleers, Drieghe, & Duyck, 2015), although it is mediated by some factors, such as age and other individual differences (Whitford & Titone, 2017; see also Wolter & Yamashita, 2018, for collocations).

THE SLA OF VARIABLE STRUCTURES

Research on the L2 acquisition of variable structures covers many languages and a range of grammatical properties, extending from phonetic to morphosyntactic to lexico-semantic (see Geeslin & Long, 2014, and De Vogelaer & Katerbow, 2017, for overviews). For example, while research on English (e.g., Adamson & Regan, 1991; Major, 2004) and French (e.g., Regan, Howard, & Lemée, 2009; Rehner, Mougeon, & Nadasdi, 2003) once dominated, we now see examples of research on variation in Spanish (e.g., Kanwit, 2017), Norwegian (Karlsen, Geva, & Lyster, 2016), Arabic (Raish, 2015), and Mandarin (Li, 2014), to name a few. In general, this field explores the linguistic, social, and individual factors that predict patterns of use and is characterized by complex statistical modeling that details the relative importance of these many factors. In general, for L2 learners at least, research demonstrates that the linguistic factors have greater influence than the social characteristics of the speaker, the hearer, or the interactional context.

Within this field of study, some work addresses the connection between the properties of individual lexical items (e.g., lexical frequency) and variable patterns of use. As with the usage-based studies, the underlying question is whether patterns of lexical diffusion, which reveal that change (or acquisition) does not affect all words at the same time or to the same degree, are attested among L2 learners. Building on work by Erker and Guy (2012), who showed the effect of lexical frequency on the patterns of use of subject forms by NSs of Spanish, Linford and colleagues explored both the way lexical frequency was measured (within corpus vs. across corpora and native vs. nonnative datasets) and the role that this factor played in the patterns of use for adult bilinguals, both Spanish and English dominant (Linford & Shin, 2013; Linford, Long, Solon, & Geeslin, 2016). The role of frequency differs across studies. Frequency may play a role as a main effect in predicting subject pronoun use (Bayley, Holland, & Ware, 2013), it may mediate the effect of other variables (e.g., revealing significant differences within tense-mood-aspect and semantic class when those variables are divided according to frequency, as in Erker & Guy, 2012; Linford & Shin, 2013; Linford et al., 2016), or it may reveal neither effect (e.g., in native Mandarin subject expression in Li & Bayley, 2018). Likewise, Solon, Linford, and Geeslin (2018) examined the phenomenon of /d/-deletion in intervocalic contexts in Spanish among native and advanced nonnative speakers and showed that the role of frequency was different for these two groups and, in fact, that frequency had an even greater effect on the nonnatives than the NSs. In light of these diverse findings, we view frequency as a useful tool, rather than a single-factor explanation.

RESEARCH ON INTERPRETATION

Research on L2 learners has further shown that differences may exist between the language produced by learners and their own patterns of interpretation. Furthermore, studies of interpretation in L2s across multiple theoretical approaches show that linguistic and individual characteristics influence patterns of processing, and that these patterns of influence change over the course of development (e.g., Borgonovo, Bruhn de Garavito, & Prévost, 2015; Cheng & Almor, 2017). Additionally, L2 processing and interpretation patterns are often different from those of NSs (e.g., Clahsen & Felser, 2006; Grüter, Lew-Williams, & Fernald, 2012). As would be expected, given what is

known about variable structures in L2s, interpretation may further be affected by implicit language attitudes, by individual identities, and by experiences and contact with particular speech communities (Schmidt, 2018; Schoonmaker-Gates, 2017). It is important to note that the term *interpretation* may be used to refer to online processing as well as more deliberate interpretation of meaning (or implied meaning), both linguistic and social, and this range of constructs further contributes to the variety of studies included under the umbrella term. Henceforth, we narrow our scope to include studies of interpretation of variable structures conducted within variationist or usage-based frameworks, as these most closely inform the current project.

Within the variationist framework, we have seen that individual lexical items do play a role in interpretation, and this role varies across levels of experience. One set of studies that has explored this issue for L2 Spanish focuses on the interpretation of verb forms in subjunctive and indicative moods in adverbial clauses. Prescriptively, for a given set of adverbs, verb forms in the subjunctive mood in adverbial clauses are predicted to have a "not yet completed" interpretation whereas those in the indicative are predicted to be interpreted as habitual events that are regularly occurring. Kanwit and Geeslin (2014) used a written interpretation task that asked participants to indicate whether an event was habitual, had not yet occurred, or if both interpretations were possible. They examined the role of the verb form (subjunctive or indicative mood), the position of the main clause relative to the adverbial clause, the regularity of the verb form, and the adverb, showing that lower-level learners did not use mood to interpret the utterances. Compared to NSs, the learners in general demonstrated less use of lexically specific patterns of interpretation and less distinction across different adverbs (*cuando* "when," *hasta que* "until," *después de que* "after"). To more effectively investigate the role of individual adverbs and their patterns of co-occurrence, Kanwit and Geeslin (2018) created a new written instrument that manipulated three of the same factors (verbal mood, verbal morphological regularity, and the adverbial conjunction), but they increased the number of adverbs included from three to six and classified these adverbs according to their degree of variability and their frequency of occurrence, based on corpus data. Their analysis showed the influence of verbal mood from the intermediate level upward, as well as a high degree of variability based on individual lexical items and their patterns of co-occurrence with subjunctive and indicative moods. From these two studies, we see that learners are able to move toward nativelike patterns of interpretation. We further see that the particular variable structure, its linguistic context, and the type of variation attested in native speech all influence the path of L2 acquisition. These findings signal the need for a careful look at these same factors in other variable structures.

THE COPULA CONTRAST IN SPANISH

To further the study of the relationship between patterns of use, including lexical frequency, and the acquisition of the ability to interpret meaning from variable structures in an L2, we have selected the copula contrast in attributive contexts in Spanish as the variable structure for examination. In what follows, we provide a brief overview of how this structure operates in native Spanish, including attested patterns of variation and an account of what is known to date about the role that particular lexical items (i.e., adjectives) play in the patterns of use. We continue with an overview of what has been discovered for the L2 acquisition of this structure while highlighting the attributive context.

A copular verb is a connecting element and is attested in a variety of forms in the world's languages. In Spanish, the verbs *ser* and *estar* both mean "to be" and are used to connect a referent with an attributive adjective. This is illustrated in examples (1a–b) in the following text. In this same context, there are other verbs, such as *parecer* "to seem" and *sentirse* "to feel," that carry additional lexical meaning, which can also fulfill the role of connecting a referent with an attribute (1c–d).

(1a) *Ella es alegre.*

She is a happy person.

(1b) *Ella está alegre.*

She is happy/in a good mood/looks happy.

(1c) *Ella parece inteligente.*

She seems intelligent.

(1d) *Ella se siente triste.*

She feels sad.

In addition to extensive theoretical research, which has attempted to define the role of these verbs and to delimit the contexts in which each appears (e.g., Camacho, 2012), the field enjoys substantial empirical work on the variation attested for the copula contrast. Most studies focus on *ser* and *estar* (leaving verbs like *parecer* aside) and define the context of analysis by the function of attribution, although these contexts can be defined syntactically as the [copula + adjective] context as well (e.g., Brown & Cortés-Torres, 2012; Díaz-Campos & Geeslin, 2011; Geeslin, 2014; Geeslin & Guijarro-Fuentes, 2008; Gutiérrez, 2003). Across studies, periods, and geographic regions, the most generalizable result is that Spanish is undergoing a process of language change through which the use of *estar* is extending into contexts where *ser* was previously the only form attested. In other words, contexts that were formerly categorical are now variable, allowing both copulas. The selection between these two forms has been associated with a range of factors, including social characteristics, geography, elicitation task, as well as linguistic factors, such as predicate type, experience with the referent, adjective class, frame of reference, and the referent's animacy and changeability (see, for example, Díaz-Campos & Geeslin, 2011, for Caracas; Gutiérrez, 2003, for Michoacán, Mexico, and Houston; Brown & Cortés-Torres, 2012, for Puerto Rico; Malaver, 2012, for Mexico and Guatemala; Geeslin & Guijarro-Fuentes, 2008, for varieties of Peninsular Spanish; and Silva-Corvalán, 1986, for Los Angeles). In the review that follows, we highlight studies that explore the relationship of individual adjectives to patterns of use.

In an analysis of patterns of co-occurrence between individual adjectives and the copulas in Spanish, Geeslin (2014) found that in her U.S.-based NS corpus of 764 examples of [copula + adjective] structures, only 9 of 26 adjectives that were used five or

more times occurred with both *ser* and *estar*. In the face of earlier work that suggested that up to 80 percent of all adjectives allowed variation between *ser* and *estar* (e.g., Mesa Alonso, Domínguez Herrera, Padrón Sánchez, & Morales Aguilera, 1993), this was somewhat surprising. In related work, Brown and Cortés-Torres (2012), in an analysis of 2,566 tokens extracted from informal conversations in Puerto Rican Spanish, argued that [*ser* + adjective] and [*estar* + adjective] are constructions and that specific [copula + adjective] constructions have definitive pragmatic and semantic uses. Their analysis showed that *estar* occurs with an individual frame of reference, mental and physical state adjectives, resultant states, and immediate experience with the referent. However, there were 116 additional uses of [*estar* + adjective] that were not explained by these factors. A closer analysis of these cases showed that more than half of the tokens (*n* = 60) were produced using just seven adjectives, such as *brutal* "excellent," *bueno* "good," and *bonito* "pretty." They argue that these adjectives, which have high token frequency, are derived from analogical extensions to speakers' previous experience with the construction. They then act as prototypes for other similar adjectives. Another important finding is that individual adjectives showed different rates of use with *estar*. For example, *fuerte* "strong," *grande* "big," *difícil* "difficult," and *nuevo* "new" occurred with both copulas, but less than 25 percent of their occurrences were with *estar*. However, *loco* "crazy" and *malo* "bad" occurred with both verbs but with *estar* 51–75% of the time, while *brutal* "excellent" and *seguro* "sure" were used with *estar* more than 75% of the time. Together these studies suggest that variability may be a function of the individual lexical item as well as the semantic class of these adjectives (e.g., prototypical meanings) and each of these factors may interact with the overall lexical frequency of the adjective to influence patterns of use.

SLA RESEARCH ON COPULA CONTRAST

The body of L2 research addressing this grammatical structure is robust. Early research focused on all functions of the copula, and generally sought to confirm, refute, and refine the attested stages of acquisition across learners and learning contexts (e.g., Ryan & Lafford, 1992). The Spanish copula contrast has further been examined to advance theoretical perspectives (e.g., Bruhn de Garavito & Valenzuela, 2008; Perpiñán, Marín, & Moreno Villamar, 2019), pedagogical approaches (Cheng, 2002; VanPatten, 2010), and inquiry into understudied language pairs (e.g., Korean [Geeslin & Long, 2015]; Dutch [Pinto & Guerra Rivera, 2015]), the role of tasks in research on learner language (Geeslin, 2006), and L2 processing (Dussias, Contemori, & Román, 2014). Contemporary research has also isolated particular copular functions (e.g., locative [Dussias et al., 2014]; stative and eventive passives [Bruhn de Garavito & Valenzuela, 2008]). Consistent with this development, many studies limit their analysis to the attributive function (i.e., the [copula + adjective] context) (e.g., Geeslin, 2003).

Most studies that focus on the attributive function recognize the variability that exists in this context. This work has revealed that learners do not limit themselves to only *ser* and *estar* in these attributive contexts, but rather, they produce a range of forms, including *parecer* "to seem," *sentirse* "to feel," and *verse* "to look," just as NSs do (Geeslin & Gudmestad, 2010). Additionally, at lower levels of proficiency, attributive uses of *buscar* "to look" and *mirar* "to watch" are also attested for English-speaking learners (Geeslin, 2000) as are *hacer* "to do" and *poner* "to place" for Korean-speaking learners

of Spanish (Geeslin & Long, 2015). What is particularly interesting about these results when considered together is that they suggest a role for each lexical item and, further, that some of these may reflect L1 influences while others appear to be attested in native and nonnative language patterns, regardless of learner backgrounds.

This research demonstrates changes in the distribution of these forms across the time course of L2 acquisition. The general finding that *ser* is initially overgeneralized and then *estar* is gradually integrated into the learner grammar (e.g., Ryan & Lafford, 1992) holds true in the attributive context (e.g., see Cheng, Lu, & Giannakouros, 2008, for L1 Chinese; Geeslin, 2000, for L1 English; and Geeslin & Long, 2015, for L1 Korean). In each of those studies, *estar* use increases with proficiency, gradually replacing *ser* in the contexts where the latter had been overgeneralized. In most cases, the more proficient learners approximate nativelike rates of use of *estar* (e.g., Geeslin, 2003), but there are examples of highly advanced learners yielding significantly different rates of use or selection of *estar* (see Geeslin & Guijarro-Fuentes, 2006, for Portuguese-speaking learners who significantly exceed the native norms). Given the varying elicitation tasks and L1 backgrounds across studies, it is difficult to pinpoint the source of the differences without further investigation.

Finally, the studies in the variationist research paradigm have identified several independent factors that generally predict the patterns of copula use for L2 learners. Firstly, the influence of linguistic factors appears to outweigh that of social, interactional, or individual background factors. Thus far, factors such as learning context (e.g., study abroad vs. at-home classroom) or speaker gender do not appear to hold the same predictive power that the morphosyntactic, semantic, or pragmatic factors do. A second common finding is that the pragmatic constraints on patterns of use override the semantic ones, even at higher levels of proficiency. For example, Geeslin (2000), in a study of four levels of English-speaking learners of Spanish, found that constraints such as whether the referent + adjective combination was susceptible to change (i.e., a semantic factor) were a significant predictor of learner language patterns in early levels, but frame of reference, which is a pragmatic constraint, was only significant at higher levels. Likewise, Geeslin (2003) showed that for very advanced learners, but not for NSs, pragmatic constraints, such as experience with the referent and frame of reference, were significant predictors of *estar*. Building on this, Cheng, Lu, and Giannakouros (2008) showed a gradual shift from formulaic use to lexically based strategies to more semantic and pragmatically driven strategies at the highest level.

The role of lexical patterns attested in corpora and their relationship to learner patterns of use was investigated in Geeslin (2013). Analyzing interview data collected from graduate-level English-speaking learners and native Spanish speakers, she identified all adjectives used more than five times and categorized them by patterns of co-occurrence. Her analysis showed that while NSs produced both *ser* and *estar* with nine of the most frequently used adjectives, nonnative speakers only produced three adjectives with both *ser* and *estar*: *amable* "friendly," *feliz* "happy," and *mejor* "better."[1] Geeslin further showed that the patterns of co-occurrence were not the result of lesser lexical diversity in the patterns of nonnative speakers, as they used approximately as many different adjectives as NSs. Instead, the author suggests that these findings indicate that learners rely more on the individual lexical item (adjective) than NSs when determining patterns of copula use.

SUMMARY: PATTERNS OF USE AND THE INTERPRETATION OF THE COPULA CONTRAST

The preceding review identifies a relative lag in research assessing the L2 acquisition of the interpretation of sociolinguistically variable structures. We know that usage-based models provide a foundation through which the role of type and token frequency in L2 interpretation can be explored, and also that this issue has been addressed with success in studies of the Spanish subjunctive. In the case of the copula contrast, research suggests that the type (e.g., semantic class of an adjective, such as physical description) and the token (i.e., the individual adjective) influence production of the [copula + adjective] structure. We further know that even adjectives within the same semantic class do not always pattern in the same way. Instead, it may be the case that L2 learners follow item-based strategies (i.e., based on individual lexical items such as adjectives), and it is further likely that other factors, both linguistic and social, are also at play. The goal of the present study is to connect these findings to explore the interpretation of the copulas in Spanish when paired with adjectives that differ by adjective class and by attested patterns of co-occurrence.

THE CURRENT STUDY

Given the important role attested for individual lexical items and the gap in research examining the L2 interpretation of variable structures, the study was designed to answer the following questions:

1. How do learners and native speakers of Spanish interpret the copulas *ser* and *estar* with adjectives?
   a. How do interpretations change as level of experience with Spanish increases?
   b. How are interpretations influenced by the copula, grammatical gender, adjective class, and lexical item?
   c. How do interpretations differ based on individual lexical items within the co-occurrence classes?

PARTICIPANTS

We collected data from 86 English-speaking learners of Spanish at four levels of enrollment and from a group of 10 NSs of Spanish, who were living in the United States (i.e., also bilingual) at the time of the study. The native speakers were from several regions in the Spanish-speaking world and were deliberately invited because they represent the range of varieties of Spanish to which U.S.-based students are exposed and the bilingual norm that constitutes the target for our classroom learners. Their results are intended to provide a view of the range of possibilities that exist for NSs of Spanish rather than a characterization of any single variety of Spanish. Their countries of origin included Colombia (3), Spain (3), Mexico (2), Argentina, and Puerto Rico, and all had resided in the contiguous United States for at least one year. The learners were enrolled at a large, midwestern public university and were drawn from fourth-semester classes (i.e., the final course required to meet the language requirement), fifth-semester classes (i.e., the bridge for all majors and minors), advanced courses (i.e., for fourth-year students majoring in Spanish), and also from the graduate student population (i.e., students who

are also generally assigned to teach language courses as part of their degree program). As with our other learners, the graduate students were all U.S. born to monolingual English-speaking families. Whereas 0/25, 2/25, and 11/25 of the learners in the first three groups had spent more than three weeks abroad in a Spanish-speaking country,[2] all graduate students reported extensive experience (six months to 2.5 years, $M = 17.5$ months) studying or working abroad in Spain, Mexico, Argentina, Puerto Rico, Chile, or Peru. The graduate group approximated the age range of the NSs. The characteristics of each of these groups, including age and grammar test results are summarized in Table 1.

INSTRUMENTS

Each participant completed three tasks. The first was a written interpretation task, described in greater detail in the following text. The second task was a language background questionnaire designed to provide a snapshot of the learner characteristics that might be a source of difference between learners. This questionnaire asked about the collegiate courses taken thus far; the years of high school study completed by each participant; the purpose, location, and length of time spent abroad; and experience with other languages aside from English and Spanish. Finally, the third task was a discrete-point grammar test that included 22 multiple-choice items. The items assess formal knowledge of grammar topics generally covered in instructed settings. The scores on this test do not provide a comprehensive view of learner proficiency but they do serve to confirm the original placement of our learners into groups according to level of enrollment.

Written Interpretation Task

The linguistic data analyzed in the study come from a written interpretation task containing 24 items, each of which asked the participant to select the sentence ending that best matched their interpretation of the main clause, or to indicate that both were possible.[3] The response options were two clauses, one of which represented a class frame interpretation (i.e., comparing the referent to a class of other individuals) and the other that represented an individual frame interpretation (i.e., comparing the referent to himself/herself at another point in time). Each item on the instrument manipulated the adjective, which was classified by its semantic class (i.e., personal characteristic, physical description, or mental/physical state) and its pattern of co-occurrence with the copular verbs (*ser* for personal characteristics, *estar*

TABLE 1.   Description of participants

Group	Number of Participants ($n = 96$)	Age			Grammar Test Results		
		Mean	Range	SD	Mean	Range	SD
4th semester	$n = 25$	19.3	18–24	1.3	9.6	5–15	2.2
5th semester	$n = 25$	18.7	18–20	0.7	11.2	6–18	2.9
4th year	$n = 25$	20.3	19–22	0.9	14.6	9–21	3.7
Graduate students	$n = 11$	26.7	23–38	4.1	19.9	17–22	1.6
Native speakers	$n = 10$	29.8	23–39	5.7	21.5	21–22	0.5

for states, *ser* and *estar* for physical descriptions). Three personal characteristic adjectives were used (*interesante* "interesting," *pobre* "poor," and *diligente* "diligent"), along with three states (*contento* "happy," *enojado* "angry," and *embarazada* "pregnant"), and three physical descriptions (*flaco* "thin," *gordo* "fat," and *calvo* "bald"). Only felicitous combinations were included in the instrument, and adjectives were selected based on their rates of co-occurrence with the copulas in the corpus search, as described in the next section. Grammatical gender was also manipulated, creating one masculine and one feminine context for the nine adjectives included. There are two exceptions to this: we used only masculine referents with *calvo* "bald" and only feminine referents with *embarazada* "pregnant" and there are two such contexts of these physical description adjectives on the instrument so that each adjective occurs twice. This yielded a total of 24 items reflected in Table 2.

In addition to the factors that were manipulated, we controlled several other linguistic variables to avoid introducing additional sources of variability into the instrument. All items contained third-person singular referents, all referents were human, and all verb forms were in the present indicative. Markers that indicate surprise or that promote a reading that is either momentary or ongoing were avoided as these also contribute to copula interpretation. A sample item is shown in (2), along with how it was coded according to the linguistic variables manipulated in the instrument design.[4]

(2) Sample item from written interpretation task

Blanca está flaca ...
   a. ____ y siempre ha querido ser modelo.
   b. ____ y los tíos creen que puede ser por el estrés.
   c. Las dos son posibles.

"Blanca looks thin ...
   a. ____ and she has always wanted to be a model.
   b. ____ and our aunt and uncle think it could be due to stress.
   c. Both are possible."

Coding: [physical description adjective], [feminine gender], [*estar*].

For all items in the instrument, the prescriptive prediction is that if *estar* occurs in the main clause, the referent is presented in an individual frame and if *ser* occurs in the main clause the interpretation should correspond to a class frame. However, based on the previous literature on variation and the copula contrast, we anticipate that this trend is not categorical. We further expect that differences in patterns of interpretation may be

TABLE 2.   Design of items in written interpretation task

Factors	Total number of items
3 personal characteristic adjectives with *ser* × 2 genders	6
3 mental/physical state adjectives with *estar* × 2 genders	6
3 physical description adjectives × 2 genders × 2 copulas (*ser* and *estar*)	12

attested at the level of the individual adjective and also that the grammatical gender factor will hold greater influence at lower levels of proficiency because of the attested tendency to overgeneralize masculine forms at early stages of development (see Alarcón, 2014 for overview).

Corpus Search to Select Adjectives for the Instrument

To represent diverse semantic classes and to select adjectives within each class that empirically occur nearly exclusively with *ser*, with *estar*, or that occur with both, we searched the two-billion-word web/dialect version of the *Corpus del español* (Davies, 2015–2016) for tokens of the verb forms *es* (from *ser*) and *está* (from *estar*) with 50 different adjectives, noting which occurred nearly exclusively with only one copula and which were well attested with both (Table 3).[5] We identified adjectives belonging to the semantic class "personal characteristic" that occurred nearly exclusively with *ser*, adjectives in the semantic class "mental/physical state" that occurred nearly exclusively with *estar,* and adjectives in the class "physical description" that exhibited variability in co-occurrence with *ser* and *estar*. We further ensured that these adjectives are relatively frequent and are known to learners even at relative lower levels of enrollment. There are, of course, adjectives in these semantic classes that may exhibit different patterns and, thus, these preliminary steps are important in making inferences from our results about the role of individual lexical items and general classifications based on patterns of co-occurrence.

RESULTS

We present the overall rates of selection for our interpretation response options before turning our attention to mixed-effects regression models. The first model includes all task items. The next two analyze the subset of items that conveyed physical descriptions through both copulas.

TABLE 3.   Partial summary of search of *es/está* + ADJ in *Corpus del español* (Davies, 2015–2016 two-billion-word web corpus)

Adjective	Type	*ser* count	*ser* %	*estar* count	*estar* %
interesante "interesting"	personal characteristic	21,053	**98.58**	304	1.42
pobre "poor"	personal characteristic	3,150	**98.68**	42	1.32
diligente "diligent"	personal characteristic	76	**100.0**	0	0.00
contento "happy"	mental/physical states	10	0.54	1,851	**99.46**
enojado "angry"	mental/physical states	6	1.15	518	**98.85**
embarazada "pregnant"	mental/physical states	9	0.20	4,501	**99.80**
calvo "bald"	physical description	109	**85.16**	19	**14.84**
flaco "thin"	physical description	94	**75.81**	30	**24.19**
gordo "fat"	physical description	241	**66.94**	119	**33.06**

Overall Rates of Interpretation

For NSs, the class response was the most popular followed by the "both" response. This was the case for the graduates and fourth-year learners as well (Table 4). Although the class frame was the most common response for the fourth- and fifth-semester learners, the individual frame was a more frequent interpretation than the "both" response.

Mixed-Effect Models: All Items

In all regression models we present, we have combined the individual frame interpretation with the response that both interpretations were possible because (a) increased acceptance of the individual interpretation (i.e., the interpretation traditionally associated with *estar*) is in the direction of language change (i.e., extension of *estar*) and (b) the individual and "both" responses patterned similarly in cross-tabulations, unlike the class response, indicating that participants treated the former two responses differently from the latter.[6] For each participant group, we first ran a mixed-effects model in Rbrul (Johnson, 2009) that considered all items from the task. In this model, the two applicable independent linguistic variables were adjective class and grammatical gender, which were run as main effects, and the individual participant and the adjective were run as random effects. In all models, acceptance of the individual frame was entered as the input value, meaning that factor weights closer to 1 indicate a favoring effect for the individual frame, whereas values closer to 0 indicate a disfavoring effect (i.e., favoring of the class frame). The range is provided for significant variables and is the difference between the factors that most and least favored the individual frame response. *P*-values for each variable indicate whether the model selected that variable as significant (i.e., for values < .05) and a correspondence between low *p*-values and high ranges can be noted. The overall rate of individual frame allowance can be found at the bottom of the table.

TABLE 4. Distribution of interpretation responses

Group	Individual		Class		Both		Total	
	#	%	#	%	#	%	#	%
4th Semester (*n* = 25)	203	33.8	255	42.5	142	23.7	600	100
5th Semester (*n* = 25)	172	28.7	275	45.8	153	25.5	600	100
4th Year (*n* = 25)	141	23.5	248	41.4	210	35.1	599	100
Graduates (*n* = 11)	76	28.8	100	37.9	88	33.3	264	100
NSs (*n* = 10)	50	20.8	122	50.8	68	28.3	240	100
Group total:							2,303	

Our mixed-effects models evaluate the linguistic elements that influence these distributions.

When all items from the task were considered (Table 5), adjective class was significant for NSs, with mental states (which occurred with *estar*) favoring the individual frame interpretation. For learners, adjective class was not significant for fourth-semester learners but was at each subsequent level. Although adjective class was not significant at the fourth semester, learners selected the individual interpretation less with mental states than their overall response rate, whereas at the later levels, this context favored the individual response, as it did for NSs, with factor weights increasing at each learner level until slightly overshooting the NS weight. Personal characteristics (with *ser*) disfavored the individual interpretation for NSs, which was also the case for all learner groups after the third semester, with factor weights again moving toward and eventually surpassing those of the NSs. For all participant groups, physical descriptions (which occurred half with *ser* and half with *estar*) fell between these extremes. For the NSs and fifth-semester and graduate learners, participants allowed the individual interpretation less with these items than with the average item (i.e., these slightly disfavored the individual response), whereas for fourth-year learners this context slightly favored the individual interpretation.

The grammatical gender variable, which applied to all items, was not significant for NSs or graduate learners, but did significantly constrain responses for the three lower levels. In each of those cases, feminine referents favored the individual interpretation and masculine referents disfavored the response, meaning that these groups responded that a female referent was more likely to be compared to herself than a male referent was. This variable was most predictive at the lowest level, at which adjective class was not significant.

Mixed-Effects Models: Items Divided by Copula

Because NSs and learners (after fourth semester) interpreted sentences differently based on which copula appeared, we now divide our data in half, first considering only those

TABLE 5.   Factors predicting allowance of individual frame (all items)

Groups and Factors		4th Sem.	5th Sem.	4th Year	Graduates	Native Speakers
Adjective Class	Mental States (*Estar*)	[0.432]*	0.669	0.737	0.860	0.783
	Physical Descriptions (Variable)	[0.471]	0.475	0.521	0.387	0.495
	Personal Characteristics (*Ser*)	[0.596]	0.354	0.247	0.205	0.220
	Range	–	*32*	*49*	*65*	*56*
	p-value	0.144	0.0172	0.000464	4.02e–05	0.000257
Grammatical Gender	Feminine	0.568	0.550	0.563	[0.442]	[0.564]
	Masculine	0.432	0.450	0.437	[0.558]	[0.436]
	Range	*14*	*10*	*12*	–	–
	p-value	0.00518	0.0355	0.013	0.155	0.0888
Participant	*Rand. St. Dev.*	*0.437*	*0.322*	*0.503*	*0.000*	*0.487*
Adjective	*Rand. St. Dev.*	*0.258*	*0.353*	*0.320*	*0.135*	*0.000*
Overall rate of individual frame allowance		57.5%	54.2%	58.6%	62.1%	49.2%

*Bracketed numbers indicate that the factor was not significant in the regression.

items that contained *ser* and then only those containing *estar*. This enables us to determine whether additional factors constrained interpretation when the copula is controlled and whether participants responded differently for different adjectives within the same copula category. For these subsets, the independent variables evenly represented across the items were the adjective used and grammatical gender of the referent. These two independent variables were entered into a mixed-effects model for each group, along with individual participant as a random effect.[7]

For the subset of items that included *ser* (Table 6), NSs were constrained by the individual adjective used, as were the fourth- and fifth-semester and fourth-year learners. The fourth-semester learners were also constrained by grammatical gender. For overall rates of interpretation (final row of the table), participants reduced allowance of the individual frame in these items containing *ser* across groups, beginning with 58.3% at the fourth-semester level and reaching a low of 36.7% in the NS group.

Examining the individual adjective used, the NS data revealed two interesting results that show sensitivity to adjective class, but also to the individual adjective. Firstly, the physical description adjectives (*calvo* "bald," *gordo* "fat," *flaco* "thin") roughly grouped together as those that most favored the individual frame interpretation, as opposed to the personal characteristics adjectives (*interesante* "interesting," *diligente* "diligent," *pobre* "poor") that generally disfavored this response, although *interesante* behaved similarly to *flaco* from the former group. Secondly, NSs showed difference within the two adjective classes, as, for example, *calvo* favored the individual interpretation much more than *flaco*, and *pobre* disfavored the response much more than *interesante*. Similarly, whereas *diligente* and *pobre* (personal characteristic adjectives) disfavored the individual interpretation for NSs, they both favored this response for the fourth-semester group and one of them (*pobre*) favored it for the fifth-semester learners. Adjective class was not significant for the

TABLE 6. Factors predicting allowance of individual frame for *ser* items

Groups and Factors		4th Sem.	5th Sem.	4th Year	Graduates	NSs
Adjective	*calvo* "bald"	0.304	0.586	0.777	[0.721]	0.804
	gordo "fat"	0.640	0.726	0.655	[0.336]	0.665
	flaco "thin"	0.322	0.276	0.307	[0.447]	0.505
	interesante "interesting"	0.566	0.390	0.332	[0.447]	0.505
	diligente "diligent"	0.615	0.479	0.477	[0.592]	0.379
	pobre "poor"	0.566	0.544	0.429	[0.447]	0.162
	Range	*34*	*45*	*47*	*–*	*64*
	p-value	*0.00154*	*0.000305*	*1.98e-05*	*0.191*	*0.00469*
Grammatical Gender	Feminine	0.611	[0.542]	[0.565]	[0.543]	[0.514]
	Masculine	0.389	[0.458]	[0.435]	[0.457]	[0.486]
	Range	*22*	*–*	*–*	*–*	*–*
	p-value	*0.00138*	*0.216*	*0.0654*	*0.409*	*0.816*
Participant	*Rand. St. Dev*	0.541	0.520	0.724	0.174	0.573
Rate of individual frame allowance w/ *ser* items		58.3%	43.7%	41.5%	37.9%	36.7%

graduate students, who tended to treat adjectives that occurred with *ser* more similarly to each other than the other groups did.

Although the individual adjective was significant for the first three learner levels, none of those groups behaved like the NSs, as the adjective rankings did not coincide with adjective class, with at least one physical description adjective disfavoring the individual interpretation for these groups (and two doing so for the fourth-semester learners). Grammatical gender was only significant for the least experienced learner group, for whom feminine referents favored the individual interpretation. Although grammatical gender was not significant for the other groups, the individual interpretation was selected more with female referents than male referents for the remaining groups.

For the other half of task items, which included the verb *estar*, NSs were again constrained according to the individual adjective, as were fourth- and fifth- semester learners (Table 7). NSs and fourth-year learners were also constrained by grammatical gender for these items. Overall rates of individual frame interpretation also revealed a pattern across groups, as learners increasingly allowed the individual interpretation, beginning at 57.3% at the fourth semester and ending at 86.4% for the graduates. Nevertheless, all of these rates exceeded that of the NSs beginning at the fifth semester.

Once again, and even more clearly than with the *ser* subset, NSs responded to adjectives differently according to adjective class, with the three mental state adjectives (*contento* "happy," *embarazada* "pregnant," *enojado* "angry") favoring the individual interpretation and the three physical description adjectives disfavoring the response. Also similar to the *ser* items, although adjectives patterned along these classes, NSs still demonstrated a sizable range within each class, for example favoring the individual interpretation much more with *contento* "happy" than *enojado* "angry," and disfavoring it more with *calvo* "bald" than *gordo* "fat." As with the *ser* items, learner groups did not

TABLE 7. Factors that predict allowance of individual frame for *estar* items

Groups and Factors		4th Sem.	5th Sem.	4th Year	Graduates	NSs
Adjective	*contento* "happy"	0.448	0.414	[0.600]	[0.581]	0.779
	embarazada "pregnant"	0.645	0.654	[0.524]	[0.581]	0.643
	enojado "angry"	0.346	0.636	[0.501]	[0.756]	0.516
	gordo "fat"	0.600	0.414	[0.471]	[0.287]	0.458
	flaco "thin"	0.406	0.311	[0.343]	[0.231]	0.303
	calvo "bald"	0.555	0.574	[0.565]	[0.581]	0.286
	Range	*30*	*34*	*–*	*–*	*49*
	p-value	*0.0247*	*0.00573*	*0.314*	*0.136*	*0.0389*
Grammatical Gender	Feminine	[0.500]	[0.571]	0.602	[0.500]	0.637
	Masculine	[0.500]	[0.429]	0.398	[0.500]	0.363
	Range	*–*	*–*	*20*	*–*	*28*
	p-value	*1.00*	*0.0651*	*0.0167*	*1.00*	*0.026*
Participant	*Rand. St. Dev*	0.484	0.619	0.700	1.089	0.287
Rate of individual frame allowance w/ *estar* items		57.3%	64.7%	75.7%	86.4%	61.7%

match this NS patterning for adjectives. This variable was significant at the fourth- and fifth- semester levels, although fourth-semester learners showed opposite directions of effect than the NSs (i.e., favoring instead of disfavoring of the individual interpretation and vice versa) for four of the six adjectives, and fifth semesters differed from the NSs for two directions of effect. For fourth-year learners and graduates, adjective was not significant, perhaps due to their overall high selection of the individual frame response that varied less across adjectives.

Finally, grammatical gender was significant for the NSs and fourth-year learners, with both groups favoring the individual interpretation with female referents and disfavoring it with male referents. For the NSs, this likely reflects the favoring effect that *embarazada* "pregnant" (which only appeared in the feminine) had on the individual interpretation and the disfavoring effect that held for *calvo* "bald" (only masculine in the data). For the fourth-year learners, this explanation does not apply because those two adjectives were treated similarly and the adjective was not significant for items with *estar*.

DISCUSSION

Our study was designed to answer research questions about how learners and NSs of Spanish interpreted the copulas *ser* and *estar* with attributive adjectives, along with how such interpretations changed as level of experience with Spanish increased. Results indicated that NSs selected the class frame interpretation most often overall, followed by the response that both interpretations were possible, which was the same ordering for fourth-year and graduate learners. Fourth- and fifth-semester learners shared with the other groups that the class interpretation was the most popular but selected the individual frame interpretation at higher rates than the "both" response.

Mixed-effects analyses of all items revealed that for NSs and all learner groups after the fourth semester, adjective class was the most predictive independent linguistic variable of interpretation, with mental states indicated by *estar* favoring the individual frame interpretation and physical descriptions indicated by *ser* disfavoring this interpretation. Across the learner groups, this variable became increasingly predictive, as indicated by lower *p*-values and higher ranges, with graduate learners eventually overshooting NS patterns by more strongly favoring the individual interpretation with mental states (*estar*) and disfavoring it with personal characteristics (*ser*), indicating greater prescriptivism on their part than that of NSs. This has been reported for interpretation of other variable structures (e.g., the mood contrast in Kanwit & Geeslin, 2014), and could be related to the greater proportion of formal contexts usually represented in advanced learner as opposed to NS input (Tarone, 2007).

In contrast, fourth-semester learners were not constrained by adjective class across the copulas when all items were considered, and instead had interpretation predicted by the grammatical gender of the referent, favoring the individual interpretation when referents were feminine. Although gender did not significantly affect NSs or graduate learners, it did maintain significance at the fifth-semester and fourth-year levels. To some extent,

these findings are consistent with past work on interpretation, which has shown less experienced learners to be unaffected by the variables that most influence NS interpretation and to be influenced instead by factors that do not constrain NSs. For example, Kanwit and Geeslin (2014) found that although NSs and more experienced learners were primarily influenced by mood in interpreting whether an action had already occurred, the least experienced learners were not significantly constrained by mood and instead were influenced by clause order and verbal morphological regularity, which were not significant for the other groups when all items were considered. Like that study, the current results support more general claims that in determining how to interpret a given structure, learner comprehension strategies may involve prioritizing cues that are less meaningful for NSs. For instance, chronological ordering of sentential elements occurs commonly cross-linguistically, but in determining which actions occur before others, NSs are likely to attend to temporality expressed in verbal morphology (Bardovi-Harlig, 2000; Kanwit & Geeslin, 2014, 2018), whereas learners may instead rely on the ordering of elements, assuming that information has been presented chronologically, which is especially the case at earlier, premorphological stages (i.e., pragmatic stages, Bardovi-Harlig, 2000).

Because NSs and most learners were constrained by the copula used, we ran additional analyses holding the copulas constant to determine whether participants treated individual adjectives divergently for a single copula and whether grammatical gender played a role. To answer these questions, additional mixed-effects regressions were run, which only included the half of the items containing *ser* and then the half containing *estar*. Overall rates of interpretation within each copula subset revealed development, with learners decreasing selection of the individual interpretation with *ser* at each level from the fourth semester (58.3%) to the graduate level (37.9%), who most approximated NS rates (36.7%), and increasing the interpretation with *estar* across levels, although this meant overshooting NS rates. The graduates' greater selection of the individual interpretation with *estar* provides further evidence for NS interpretations that are less prescriptive than those of advanced learners (Kanwit & Geeslin, 2014).

The individual adjective was significant in both the *ser* and *estar* regressions for the NSs, indicating that adjectives patterned together by class but within each class NSs also demonstrated lexical effects in treating individual adjectives rather distinctly in spite of the general pattern. For example, in the *estar* model, the three mental state adjectives all favored the individual interpretation, while the three physical description adjectives all disfavored it, and yet the favoring effect for *contento* "happy" was much stronger than for *enojado* "angry," with the disfavoring effect for *calvo* "bald" much stronger than that of *gordo* "fat." Thus, the NSs did treat the adjectives according to classes, but they did not respond uniformly within a given adjective class. This means that NSs showed a combination of constraint-based and lexically based interpretation. They demonstrated the former in differentiating according to adjective class when all items were considered and, when items were divided by copula, in showing results that grouped according to adjective class in both of the additional models. Nevertheless, they also demonstrated lexically based interpretation in these smaller models, as they treated each adjective divergently within a given class, showing a rather sizable range within classes. In tandem, these results support prior work that has shown NSs to use patterns of

collocation and to have a lexical basis for interpretation (Bybee, 2007, 2010; Edmonds & Gudmestad, 2014; Goldberg, 2013; Kanwit & Geeslin, 2014, 2018).

Learners did not match NSs in grouping adjectives according to type within the individual verb regressions, although they did demonstrate development in no longer being significantly constrained according to directions of effect that opposed the NSs'. For example, for the *estar* models, the fourth-semester learners had opposite directions of effect vis-à-vis the NSs for four of the six adjectives (and also for the *ser* models), whereas this was the case for only two of the adjectives for the fifth-semester group, and then this variable was not significant for fourth-year and graduate learners, although each only diverged from the NSs for one adjective's (*calvo* "bald") direction of effect.

Like the NSs, as early as the fifth semester, learners also tended toward constraint-based interpretation when all items were considered, as they differentiated according to adjective class. Unlike the NSs, graduate learners did not further differentiate within adjective classes, as individual adjective was not significant in either verb-specific model. Thus, the more experienced learners seem to have systems that are the most similar to what would be predicted by constraint-based interpretation: adjective class is highly predictive when all items are considered and individual adjective is not significant in either copula's model for the graduate learners and is not significant in the *estar* model for the fourth-year learners. The learner tendencies to interpret *estar* as expressing the individual frame and to overshoot NS norms support prior research on the copula contrast (Geeslin, 2003; Geeslin & Guijarro-Fuentes, 2006). Learners overshooting NS norms in interpretation is also consistent with work on other structures (e.g., mood interpretation in Kanwit & Geeslin, 2014, 2018). Less differentiation based on individual adjectives for graduate learners is similar to prior work that has shown NSs to rely on collocations more than learners (Edmonds & Gudmestad, 2014) and for advanced learners to depend more on class than lexical item in interpretation (e.g., significant results for both individual adverb and mood for NSs, but only for mood for graduates in Kanwit & Geeslin, 2014).

The least experienced learners are not constrained by adjective class when all items are considered and select the individual interpretation less with mental states than with personal characteristics or physical descriptions. This means that they have not yet developed constraint-based interpretation for the copula contrast. Rather than constraint-based interpretation, they respond according to the adjective used, although they do so in ways that usually diverge from the NSs (for four of the six adjectives for each copula's model). In the absence of other cues for interpretation, they also afford a larger role for grammatical gender, which shapes their interpretation when all items are considered and in the *ser* model.

The learners with intermediate experience (i.e., fifth semester and fourth year) fell between the two extremes presented. They demonstrated more constraint-based interpretation than the fourth-semester learners because adjective class achieved significance for them. Nevertheless, for them, the individual adjective is also significant, with individual adjectives increasingly behaving like they do for NSs in each verb's models, rather than yielding opposing directions of effect. Thus, the data seem to indicate that learners have to develop constraint-based interpretations (i.e., they do not yet have them at the lowest level), that such interpretations give way to lexical effects with more language experience, but that differential treatment within an adjective class may be reduced as participants behave in a more prescriptive fashion and even overshoot NS

conditioning. The fact that fourth-year learners demonstrate some of the just-described qualities of the intermediate groups and also some of what was described for the graduates in the preceding text (e.g., lack of significance for individual adjectives in the *estar* model) indicates that the proposed changes may indeed occur in the order proposed.

Although grammatical gender constrained the least proficient learners in the main model, it mostly did not constrain interpretation in the models that were divided by copula, with three exceptions. For the items with *ser*, grammatical gender was significant only for the least experienced learners (fourth semester), who showed a pattern similar to when all items were considered. For the items with *estar*, NSs were constrained by gender, but this was likely also an effect of individual adjectives because the feminine adjective *embarazada* "pregnant" highly favored the individual interpretation and the masculine adjective *calvo* "bald" highly disfavored it. Feminine adjectives also favored the individual interpretation for fourth-year learners, although the same discrepancy for *embarazada* vis-à-vis *calvo* did not hold for this group, with the two adjectives treated rather similarly (both favoring the individual interpretation).

The discussion thus far has provided a detailed account of how the interpretation of [copula + adjective] structures develops with increasing proficiency and how the predictive models of such interpretation differ across levels and from NSs. As we have argued, this research sits at the intersection of several subfields of inquiry and, consequently, has numerous implications across these fields. The study of the L2 acquisition of the copula contrast in Spanish benefits from new knowledge related to interpretation (vis-à-vis production) as well as new findings related to the role of semantic groups (i.e., adjective class) and individual lexical items. The varying importance of these factors between groups further connects these findings to broader questions about the degree to which linguistic behavior is rule governed (i.e., influenced by generalized patterns) or item based (i.e., specific to a given lexical item or collocation). Our findings are consistent with existent research, such as Durrant and Schmitt (2009), who found that learners produce frequent collocations at higher rates than NSs and less frequent collocations at comparatively lower rates. In the case of interpretation, our NNSs showed a lesser effect for individual lexical items than the NSs did. One key difference between Durrant and Schmitt and the current study is the manner in which lexical frequency was operationalized. In the former, the researchers compared relatively more and less frequent collocations to one another, whereas we ensured that all collocations on the instrument were regularly occurring (i.e., rather frequent). Despite this, our results point to comparable differences between NSs and NNSs. We hypothesize that this reflects differing ranges and depths of experience between traditional classroom L2 learners and NSs. Studies like Linford et al. (2016) suggest that group-specific corpora may be a better source for measuring lexical or collocational frequency for learners. Their study of subject form production showed that while NS patterns of use generally reflected patterns in other larger corpora, the learner patterns were less generalizable between group-specific and larger, publicly available corpora. Taken together, we hypothesize that the use of group-specific corpora can be used profitably to further study patterns of interpretation.

These findings also expand our knowledge of interpretation more broadly. Doing so helps to address the current bias in the literature toward analyses of production. In the case of L2 variable structures, this is especially important, as sociolinguistic competence

is generally understood to place greater burdens on accurate interpretation of socio-linguistically variable forms than on producing the full range of forms available in the input, some of which may not reflect the identity of an individual learner. With these implications in mind, we see that in addition to improving our knowledge of the development of a particular variable structure, the present study contributes more broadly to our understanding of L2 interpretive abilities, as well as usage-based and corpus research approaches to the patterning of language use more broadly.

CONCLUSIONS, LIMITATIONS, AND FUTURE DIRECTIONS

Our findings build on the existing work on the L2 acquisition of variable structures by expanding the relatively scarce knowledge of interpretation to the Spanish copula. They further connect this area of study to research on the roles of type and token frequency in native and nonnative patterns of use. We see evidence of differences between production and perception and also that the role of frequency may differ between NSs and learners in systematic ways. Because this is one of few studies examining the L2 interpretation of a variable structure in this way, many paths for future inquiry remain. Firstly, while the written instrument is held constant across participants and the individual was considered as a random effect, the participants likely also contribute to the variability attested. Future studies might examine individual results in greater detail, along with individuals' varied social characteristics and experiences. Likewise, our learner groups differ considerably in proficiency from one another and additional data from intermediate-level learners or longitudinal data will provide greater detail about developmental paths. We see great possibility in expansion through supplementary interpretive tasks (e.g., picture matching), additional adjectives and classes, and further variable structures, including those associated with particular geographic varieties. In sum, the present study lays the foundation for future research into the L2 acquisition of interpretation of socio-linguistically variable structures.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S0272263119000718

NOTES

[1]NSs used both copulas with *abierto*, *casado*, *difícil*, *feliz*, *grande*, *ingenuo*, *mejor*, *seguro*, and *tranquilo*.

[2]The two fifth-semester learners who had spent time abroad reported 2 months in Spain or 3 months in Mexico. Ten fourth-year learners reported 2–5 months in Spain, Mexico, or Puerto Rico, and one reported 2.5 years in Spain and Mexico. Learners in both of these groups matched most appropriately to their level colleagues on the grammar test and interpretation task. In fact, none of these participants had the highest grammar score in their level group.

[3]It is common for research on the acquisition of variable structures to allow that multiple options are acceptable (e.g., Geeslin, 2003; Geeslin & Guijarro-Fuentes, 2006, 2008; Geeslin & Long, 2014; Kanwit, 2018). The "both" option may provide a fuller picture of the nature of the variation and allows the researcher to analyze data in light of language changes in progress.

[4]See supplementary materials for more items and IRIS for the full task: https://www.iris-database.org/iris/app/home/detail?id=york%3a937043&ref=search.

[5]See supplementary materials for additional search results.

[6]An original multinomial regression confirmed that the individual and "both" responses had similar directions of effect in comparison to the class response.

[7]For these items, adjective is run as a main effect, not a random effect, to see the role played by the individual adjective manipulated within the physical description class. In the prior model, adjective was run as a random effect because each adjective could only occur within its one adjective class (i.e., individual adjectives could not occur across all adjective types).

REFERENCES

Adamson, H. D., & Regan, V. M. (1991). The acquisition of community speech norms by Asian immigrants learning English as a second language: A preliminary study. *Studies in Second Language Acquisition*, *13*, 1–22. https://doi.org/10.1017/S0272263100009694

Alarcón, I. (2014). Grammatical gender in second language Spanish. In K. L. Geeslin (Ed.), *The handbook of Spanish SLA* (pp. 202–218). New York, NY: Wiley.

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

Bayley, R., Greer, K., & Holland, C. (2013). Lexical frequency and syntactic variation: A test of a linguistic hypothesis. *University of Pennsylvania Working Papers in Linguistics*, *19*, 21–30.

Borgonovo, C., Bruhn de Garavito, J., & Prévost, P. (2015). Mood selection in relative clauses: Interfaces and variability. *Studies in Second Language Acquisition*, *37*, 33–69. https://doi.org/10.1017/S0272263114000321

Brown, E. L., & Cortés-Torres, M. (2012). Syntactic and pragmatic usage of the [*estar* + adjective] construction in Puerto Rican Spanish: ¡Está brutal! In K. L. Geeslin & M. Díaz-Campos (Eds.), *Selected proceedings of the 14th Hispanic Linguistics Symposium* (pp. 61–74). Somerville, MA: Cascadilla Proceedings Project.

Bruhn de Garavito, J. B., & Valenzuela, E. (2008). Eventive and stative passives in Spanish L2 acquisition: A matter of aspect. *Bilingualism: Language and Cognition*, *11*, 323–336. https://doi.org/10.1017/S1366728908003556

Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford, UK: Oxford University Press.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge, UK: Cambridge University Press.

Camacho, J. (2012). Ser and *estar*: The individual/stage-level distinction and aspectual predication. In J. I. Hualde, A. Olarrea, & E. O'Rourke (Eds.), *The handbook of Hispanic linguistics* (pp. 453–475). West Sussex, UK: Wiley-Blackwell.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*, 1–47.

Cheng, A. C. (2002). The effects of processing instruction on the acquisition of *ser* and *estar*. *Hispania*, *85*, 308–323.

Cheng, A. C., Lu, H. C., & Giannakouros, P. (2008). The uses of Spanish copulas by Chinese-speaking learners in a free writing task. *Bilingualism: Language and Cognition*, *11*, 301–317. https://doi.org/10.1017/S1366728908003532

Cheng, W., & Almor, A. (2017). The effect of implicit causality and consequentiality on nonnative pronoun resolution. *Applied Psycholinguistics*, *38*, 1–26. https://doi.org/10.1017/S0142716416000035

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42. https://doi.org/10.1017/S0142716406060024

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency effects in monolingual and bilingual natural reading. *Psychonomic Bulletin & Review*, *22*, 1216–1234. https://doi.org/10.3758/s13423-015-0819-2

Crossley, S. A., Subtirelu, N., & Salsbury, T. (2013). Frequency effects or context effects in second language word learning: What predicts early lexical production? *Studies in Second Language Acquisition*, *35*, 727–755. https://doi.org/10.1017/S0272263113000375

De Vogelaer, G., & Katerbow, M. (Eds.) (2017). *Acquiring sociolinguistic variation*. Amsterdam, The Netherlands: John Benjamins.

Díaz-Campos, M., & Geeslin, K. L. (2011). Copula use in the Spanish of Venezuela: Is the pattern indicative of stable variation or an ongoing change? *Spanish in Context*, *8*, 73–94.

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first-and second-language word recognition: A lexical entrenchment account. *The Quarterly Journal of Experimental Psychology*, 66, 843–863. https://doi.org/10.1080/17470218.2012.720994

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching*, 47, 157–177. https://doi.org/10.1515/iral.2009.007

Dussias, P. E., Contemori, C., & Román, P. (2014). Processing *ser* and *estar* to locate objects and events: An ERP study with L2 speakers of Spanish. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, 27, 54–86. https://doi.org/10.1075/resla.27.1.03dus

Edmonds, A., & Gudmestad, A. (2014). Your participation is greatly/highly appreciated: Amplifier collocations in L2 English. *Canadian Modern Language Review*, 70, 76–102. https://doi.org/10.3138/cmlr.1704

Ellis, N. C., & Ferreira-Junior, F. (2009). Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal*, 93, 370–385. https://doi.org/10.1111/j.1540-4781.2009.00896.x

Ellis, N. C., & Schmidt, R. (1997). Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition*, 19, 145–171.

Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, 42, 375–396. https://doi.org/10.1002/j.1545-7249.2008.tb00137.x

Erker, D., & Guy, G. R. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language*, 88, 526–557.

Geeslin, K. L. (2000). A new approach to the second language acquisition of copula choice in Spanish. In R. Leow & C. Sanz (Eds.), *Spanish applied linguistics at the turn of the millennium: Papers from the 1999 Conference on the L1 & L2 Acquisition of Spanish and Portuguese* (pp. 50–66). Somerville, MA: Cascadilla Press.

Geeslin, K. L. (2003). A comparison of copula choice: Native Spanish speakers and advanced learners. *Language Learning*, 53, 703–764. https://doi.org/10.1046/j.1467-9922.2003.00240.x

Geeslin, K. L. (2006). Task design, discourse context and variation in second language data elicitation. In C. A. Klee & T. L. Face (Eds.), *Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 74–85). Somerville, MA: Cascadilla Proceedings Project.

Geeslin, K. L. (2013). Future directions in the acquisition of variable structures: The role of individual lexical items in second language Spanish. In C. Howe, M.L. Quesada, & S. Blackwell (Eds.), *Selected proceedings of the 15th Hispanic Linguistics Symposium* (pp. 187–204). Somerville, MA: Cascadilla Press.

Geeslin, K. L. (2014). An analysis of co-occurrence of the Spanish copulas with adjectives: How variable are they really? In C.-T. J. Huang & F.-H. Liu (Eds.), *Peaches and plums* (pp. 419–441). Taipei: Academia Sinica.

Geeslin, K. L., & Gudmestad, A. (2010). An exploration of the range and frequency of occurrence of forms in potentially variable structures in second-language Spanish. *Studies in Second Language Acquisition*, 32, 433–463. https://doi.org/10.1017/S0272263110000033

Geeslin, K. L., & Guijarro-Fuentes, P. (2006). Second language acquisition of variable structures in Spanish by Portuguese speakers. *Language Learning*, 56, 53–107. https://doi.org/10.1111/j.0023-8333.2006.00342.x

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. https://doi.org/10.1017/S1366728908003593

Geeslin, K. L., & Long, A. Y. (2014). *Sociolinguistics and second language acquisition: Learning to use language in context*. New York, NY: Routledge.

Geeslin, K. L., & Long, A. Y. (2015). The development and use of the Spanish copula with adjectives by Korean-speaking learners. In I. Pérez-Jiménez, M. Leonetti, & S. Gumiel-Molina (Eds.), *New perspectives on the study of* ser *and* estar (pp. 293–324). Amsterdam, The Netherlands: John Benjamins.

Goldberg, A. E. (2013). Constructionist approaches. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 15–31). New York, NY: Oxford University Press. doi: 10.1093/oxfordhb/9780195396683.001.0001

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28, 191–215. https://doi.org/10.1177/0267658312437990

Gutiérrez, M. J. (2003). Simplification and innovation in US Spanish. *Multilingua*, 22, 169–184. https://doi.org/10.1515/mult.2003.009

Hoey, M. (1991). *Patterns of lexis in text*. Oxford, UK: Oxford University Press.

Johnson, D. E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass*, *3*, 359–383. https://doi.org/10.1111/j.1749-818X.2008.00108.x

Kanwit, M. (2017). What we gain by combining variationist and concept-oriented approaches: The case of acquiring Spanish future-time expression. *Language Learning*, *67*, 461–498. https://doi.org/10.1111/lang.12234

Kanwit, M. (2018). Variation in second language Spanish. In K. L. Geeslin (Ed.), *The Cambridge handbook of Spanish linguistics* (pp. 716–736). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/9781316779194.033

Kanwit, M., & Geeslin, K. L. (2014). The interpretation of Spanish subjunctive and indicative forms in adverbial clauses: A cross-sectional study. *Studies in Second Language Acquisition*, *36*, 487–533. https://doi.org/10.1017/S0272263114000126

Kanwit, M., & Geeslin, K. L. (2018). Exploring lexical effects in second language interpretation: The case of mood in Spanish adverbial clauses. *Studies in Second Language Acquisition*, *40,* 579–603. https://doi.org/10.1017/S0272263117000262

Karlsen, J., Geva, E., & Lyster, S. A. (2016). Cognitive, linguistic, and contextual factors in Norwegian second language learner's narrative production. *Applied Psycholinguistics*, *37*, 1117–1145. https://doi.org/10.1017/S014271641500051X

Li, X. (2014). Variation in subject pronominal expression in L2 Chinese. *Studies in Second Language Acquisition*, *36*, 39–68. https://doi.org/10.1017/S0272263113000466

Li, X., & Bayley, R. (2018). Lexical frequency and syntactic variation: Subject pronoun use in Mandarin Chinese. *Asia-Pacific Language Variation*, *4*, 135–160.

Linford, B., & Shin, N. L. (2013). Lexical frequency effects on L2 Spanish subject pronoun expression. In J. Cabrelli Amaro, G. Lord, A. de Prada Pérez, & J. E. Aaron (Eds.), *Selected proceedings of the 16th Hispanic Linguistics Symposium* (pp. 175–189). Somerville, MA: Cascadilla Proceedings Project.

Linford, B., Long, A., Solon, M., & Geeslin, K. (2016). Measuring lexical frequency: Comparison groups and subject expression in L2 Spanish. In L. Ortega, A. E. Tyler, H. I. Park, & M. Uno (Eds.), *The usage-based study of language learning and multilingualism* (pp. 137–154). Washington, DC: Georgetown University Press.

Major, R. C. (2004). Gender and stylistic variation in second language phonology. *Language Variation and Change*, *16*, 169–188. https://doi.org/10.1017/S0954394504163059

Malaver, I. (2012). "Estamos muy grandes ya": Adjetivos de edad con *ser y estar* en el español de México y Guatemala. *Lexis*, *36*, 191–224.

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

Perpiñán, S., Marín, R., & Moreno Villamar, I. (2019). The role of aspect in the acquisition of ser and estar in locative contexts by English-speaking learners of Spanish. *Language Acquisition*, *27*, 35–67. https://doi.org/10.1080/10489223.2019.1610408.

Pinto, M., & Guerra Rivera, A. (2015). Copula choice in adjectival constructions in Dutch L1 Spanish L2. In T. Judy & S. Perpiñán (Eds.), *The acquisition of Spanish in understudied language pairings* (pp. 309–328). Amsterdam, The Netherlands: John Benjamins.

Raish, M. (2015). The acquisition of an Egyptian phonological variant by U.S. students in Cairo. *Foreign Language Annals*, *48*, 267–283.

Regan, V., Howard, M., & Lemée, I. (2009). *The acquisition of sociolinguistic competence in a study abroad context*. Bristol, UK: Multilingual Matters.

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. https://doi.org/10.1017/S0272263103000056

Ryan, J. M., & Lafford, B. A. (1992). Acquisition of lexical meaning in a study abroad environment: *Ser* and *estar* and the Granada experience. *Hispania*, *75*, 714–722.

Schmidt, L. B. (2018). L2 development of perceptual categorization of dialectal sounds: A study in Spanish. *Studies in Second Language Acquisition*, *40*, 857–882. https://doi.org/10.1017/S0272263118000116

Schoonmaker-Gates, E. (2017). Regional variation in the language classroom and beyond: Mapping learners' developing dialectal competence. *Foreign Language Annals*, *50*, 177–194.

Silva-Corvalán, C. (1986). Bilingualism and language change: The extension of *estar* in Los Angeles Spanish. *Language*, *62*, 587–608. https://doi.org/10.1353/lan.1986.0023

Solon, M., Linford, B., & Geeslin, K. L. (2018). Acquisition of sociophonetic variation. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, *31*, 309–344.

Sun, D. (2014). From communicative competence to interactional competence: A new outlook to the teaching of spoken English. *Journal of Language Teaching and Research*, *5*, 1062–1071. https://doi.org/10.4304/jltr.5.5.1062-1070

Tarone, E. (2007). Sociolinguistic approaches to second language acquisition research—1997–2007. *The Modern Language Journal*, *91*, 837–848. https://doi.org/10.1111/j.1540-4781.2007.00672.x

van Compernolle, R. A., & Williams, L. (2012). Reconceptualizing sociolinguistic competence as mediated action: Identity, meaning-making, agency. *The Modern Language Journal*, *96*, 234–250. https://doi.org/10.1111/j.1540-4781.2012.01334.x

VanPatten, B. (2010). Some verbs are more perfect than others: Why learners have difficulty with *ser* and *estar* and what it means for instruction. *Hispania*, *93*, 29–38.

Whitford, V., & Titone, D. (2017). The effects of word frequency and word predictability during first- and second-language paragraph reading in bilingual older and younger adults. *Psychology and Aging*, *32*, 158–177. http://dx.doi.org/10.1037/pag0000151

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition*, *35*, 451–482. https://doi.org/10.1017/S0272263113000107

Wolter, B., & Yamashita, J. (2018). Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocational processing: What accounts for L2 performance? *Studies in Second Language Acquisition*, *40*, 395–416. https://doi.org/10.1017/S0272263117000237

Wulff, S., Ellis, N. C., Römer, U., Bardovi-Harlig, K., & Leblanc, C. J. (2009). The acquisition of tense-aspect: Converging evidence from corpora and telicity ratings. *The Modern Language Journal*, *93*, 354–369. https://doi.org/10.1111/j.1540-4781.2009.00895.x

Yi, W. (2018). Statistical sensitivity, cognitive aptitudes, and processing of collocations. *Studies in Second Language Acquisition*, *40*, 831–856. https://doi.org/10.1017/S0272263118000141