Article

¿*(Está/Es) Difícil*?: Variable Use of *Ser* and *Estar* by Heritage Learners of Spanish

Jamelyn Wheeler [1],[*] , Matthew Pollock [2],[*] and Manuel Díaz-Campos [1]

1 Department of Spanish & Portuguese, Indiana University, Bloomington, IN 47405, USA; mdiazcam@indiana.edu
2 Department of English & Foreign Languages, Louisiana State University Shreveport, Shreveport, LA 71115, USA
* Correspondence: jamewhee@indiana.edu (J.W.); matthew.pollock@lsus.edu (M.P.)

Abstract: The current study examines variation in copula selection in Spanish by looking at the written productions of three groups of language learners in the United States, including heritage learners, those with English as an L1, and international students with English as an L2. Research on copula variation in Spanish has pinpointed several key linguistic and social factors that influence selection; this study aims to apply these findings to heritage learners in order to determine how their acquisition differs from that of non-native language learners. This analysis used the COWS-L2H corpus of Spanish from the University of California, Davis. Examining over 8000 tokens of [adjective + copula] constructions in variable contexts where both *ser* and *estar* were used, the study tracks how linguistic and extralinguistic factors condition copula selection within the three learner groups and how these results compare to previous findings. Seven factors were predictive of copula selection: resultant state, frame of reference, adjective class, experience with study abroad, essay prompt, student age, and course level. Heritage learner copula use was found to be governed by a different set of predictors than that of learners, hinting at the variable motivations and backgrounds that influence use and reflect the identity goals of these speakers.

Keywords: heritage learners; *ser*; *estar*; copula; variable use; sociolinguistics; morphosyntax

1. Introduction

This study investigates variation in the selection of the copula verbs *ser* and *estar* ('to be') by three groups of learners, including heritage learners, L1 English speakers, and international L2 speakers of English. The results provide insight into the difference in copula selection between L2 English and heritage language learners, a phenomenon that has been documented across the Spanish-speaking world (e.g., *Ella está casada con mi padre* 'She is married to my father' versus *Ella es casada con un famoso rapero* 'She is married to a famous rapper'). This study finds that copula use differs in part based on the exposure that learners have to these forms; while classroom learners are exposed to generalized cues that affect copula selection, heritage learners may be more likely to rely upon their own innate judgment, which improves over time as they acquire a more precise awareness of copula use contexts.

In light of the rich variation that exists in copula usage, not only among learners but also among native speakers, these verbs have been the focus of sociolinguistic studies in the Spanish varieties of the United States (Silva-Corvalán 1994), Cuba (Díaz-Campos et al. 2017), Venezuela (Díaz-Campos and Geeslin 2011), Costa Rica (Aguilar-Sánchez 2009), Mexico (de Jonge 1993; Gutiérrez 1992; Juárez-Cummings 2014), Puerto Rico (Brown and Cortés-Torres 2012), and Spain (Geeslin and Guijarro-Fuentes 2006; Isasa 2014). In addition, this phenomenon has been examined in studies on language contact (e.g., Geeslin and Guijarro-Fuentes 2008). Overall, previous research has shown that, in variable contexts of copula selection, factors influencing variant choice include predicate type, resultant state, adjective class, frame of reference, susceptibility to change, and experience with the referent.

At the same time, there is a growing body of research examining the speech of heritage learners to determine how bilingual tendencies manifest among second and third generation speakers of the language as a means of tracking acquisition and identity construction among the sizable U.S. Hispanic population (e.g., Cuza and Pérez-Tattam 2016; Cuza et al. 2020; Ducar 2012; Lynch and Avineri 2021; Lynch 2012; Putnam and Sánchez 2013; Silva-Corvalán 1986, 1994, 2014). These heritage speakers are framed by Valdés (2001) as any individual who grew up in a household where a "heritage" language other than the dominant community norm was spoken, leading to a degree of bilingual abilities in both the dominant and heritage languages. This is consistent with Montrul's (2008, 2016) definition of heritage speakers, who she defines as students raised in a non-English household who may use or solely understand the heritage language while also being bilingual in English. These speakers present an interesting context for linguistic examination, particularly in the North American university system, given the power imbalance between Spanish and English in the United States (Lynch and Avineri 2021) as well as the difference in acquisitional approaches of heritage and L2 learners in the language classroom (e.g., Silva-Corvalán 2014). Using a sociolinguistic framework, the current article contributes to the understanding of the variable use of *ser* and *estar* and the differences in linguistic patterns produced by heritage speakers and L2 learners.

This study draws on sociolinguistic and acquisition research to determine whether L1 English learners of Spanish differ from L1 non-English learners of Spanish and Spanish heritage learners in their written production of *ser* and *estar*. These data are drawn from a college-level instructional community in Davis, California, using the publicly available COWS-L2H corpus of university student essays. In this project, patterns between the aforementioned groups are compared with previous research on native varieties of Spanish to determine how norms of heritage learner speech parallel both English and Spanish acquisitional patterns. Following a coding protocol described by Díaz-Campos and Geeslin (2011) for tracking the use of the Spanish copula, this article determines the extent to which heritage speakers differ from other groups. Based on previous research on bilingual speech (e.g., Díaz-Campos et al.), it is hypothesized that heritage speakers evince tendencies closer to L1 speakers of Spanish, while L2 Spanish learners are more likely to favor the copula *ser* across all contexts.

2. The Spanish Copula

Descriptive linguistic literature concerned with copula use makes a semantic distinction between the use of *ser* and *estar* in the [copula + adjective] context. Gili Gaya (1961) argues that while *ser* indicates permanent properties, *estar* conveys contingent or circumstantial states. For example, in the sentence *Pedro es saludable* ('Peter is healthy'), the attribute 'healthy' is presented as a permanent property of Peter. In contrast, in the sentence *Pedro está resfriado* ('Peter is sick'), the attribute 'sick' is treated as temporary with the use of *estar.*

Gili Gaya (1961) makes several additional distinctions to show how the use of *estar* in particular differs from that of *ser*. In a reference to Hanssen's (1913) earlier work with copulas, Gili Gaya states that the distinction between *ser* and *estar* can be seen as an aspectual consideration between perfective and imperfective predicates. In the example *El jarro está roto* ('The jar is broken'), the quality of "being broken" is the result of an action, thereby requiring the use of *estar,* while the same does not hold for *ser.* Additionally, another important consideration involves the speaker's immediate experience with the referent. For example, the phrase *Aquella nieve está fría* 'That snow is cold' conveys an immediate experience with the referent.

While earlier work such as that by Gili Gaya (1961) focused on individual factors, more recent sociolinguistic studies have shown that it is actually a constellation of differing factors that allow for an understanding of variable copula selection (Brown and Cortés-Torres 2012; Díaz-Campos and Geeslin 2011). The present study adopts a variationist sociolinguistic framework based on previous work by Geeslin and colleagues (e.g., Geeslin 2003, 2005; Geeslin and Guijarro-Fuentes 2006, 2008; Kanwit and Geeslin 2020) to better understand variable use among heritage speakers and L2 learners.

2.1. Linguistic Factors Conditioning Copula ()Variation

Brown and Cortés-Torres (2012), Díaz-Campos and Geeslin (2011), and Silva-Corvalán (2014), among others, have encountered linguistic predictors that constrain Spanish copula choice. Six of these predictors are explained in greater depth below: resultant state, adjective class, predicate type, experience with the referent, susceptibility to change, and frame of reference.

First, resultant state differentiates between adjectives that are the result of an accomplishment and those that are not. This is a semantic factor. For example, in the sentence *Estoy muy abierto a lo que la gente tiene que decir* ('I'm very open to what people have to say'), *abierto* 'open' is an attribute paired with *estar*, illustrating an adjective yielded from the verb *abrir* 'to open', which is a resultant state. Previous studies have shown that there is a tendency for resultant state adjectives to be paired with the copula *estar* (Geeslin 2003, 2005; Díaz-Campos and Geeslin 2011).

Next, a number of studies have shown that the semantic factor of adjective class predicts copula choice (Geeslin 2003, 2005; Brown and Cortés-Torres 2012; Gutiérrez 1992, 1994; Juárez-Cummings 2014; Ortíz López 2000). Adjectives associated with observable traits favor *ser,* whereas those referencing mental and physical states favor *estar.* However, it is also important to note that variable use of copulas has been described for adjectives indicating status as well as those in the context of age expression. For example, in the sentence *cuando estaba joven* ('when I was young'), *estar* is used with the adjective 'young' to indicate age.

Third, the semantic factor of predicate type plays an important role in copula selection (Aguilar-Sánchez 2009; Batllori et al. 2009; Batllori and Roca 2011; Camacho 2012; Clements 1988, 2006; Díaz-Campos and Geeslin 2011; Geeslin 2003; Juárez-Cummings 2014; Leonetti 1994; Marco and Marín 2015). This variable distinguishes between qualities characteristic of an individual and those attributed to the individual in a concrete situation. When a prediction pairs the attribute with *ser,* this distinction is termed "individual level", whereas pairing with *estar* is referred to as "stage level". An example of an individual-level predicate would be *Porque era muy chiquitico* ('Because he was very young'), as it references an individual characteristic, while an example of a stage level predicate would be *Cuando estaba chiquita. . .* ('When she was young. . .'), as it refers to a concrete situation.

The variable experience with the referent considers whether a statement is based on a direct or an indirect experience. Importantly, this pragmatic factor distinguishes a direct reaction to a particular situation. Events that denote a direct reaction from the speaker favor the use of *estar,* as in *Me gusta pero como está hoy que hace sol sí está acceptable* ('I like it, but how it is today, being sunny, is acceptable').

Fifth, susceptibility to change is a pragmatic factor that separates referents with properties that can be considered [+changeable] from those that cannot [-changeable]. In the example *El edificio era alto* ('The building was tall'), the property of the building's height is [-changeable], whereas for *Pedro estaba enfermo* ('Pedro was sick'), the illness can be seen as a [+changeable] trait. As shown in these examples, properties treated as [+changeable] favor *estar*.

Finally, frame of reference is a pragmatic factor that distinguishes between a comparison of a referent to itself at another point in time versus a comparison to a group of similar entities. A comparison of the referent with itself is predicted to favor pairing with *estar*; for example, in the sentence *Cuando estaba más joven* ('When I was younger'), the speaker makes a comparison to their own past. In contrast, the sentence *Tú eres joven y tienes una vida larga* 'You are young and have a long life' compares the referent to a general class of similar objects.

In the current study, these linguistic factors were examined in written language to compare differences in copula selection for heritage and L2 language learners of Spanish.

2.2. Extralinguistic Factors Conditioning Copula Variation

Social stratification of the use of *ser* and *estar* has been examined in several varieties of Spanish, including Venezuelan, Cuban, Puerto Rican, Costa Rican, and Mexican, as well as

among language learners. For example, in a recent study, Kanwit and Geeslin (2020) examined native and L2 speakers' sociolinguistic competence with variable structures including *ser* and *estar*. However, in contrast to the effect of the linguistic variables discussed above, the patterns associated with social variables have a greater tendency to vary according to the specific speech community being studied. Keeping this in mind, the social factors found to be significant in predicting copula use are socioeconomic status, age, level of education, and gender.

In a study on copula selection in Venezuelan Spanish, Díaz-Campos and Geeslin (2011) determined that several subcategories of speakers favor use of *estar*, including speakers with lower socioeconomic status and older Venezuelans (ages 46+). In a Mexican Spanish context, Juárez-Cummings (2014) found that members of the lower and middle classes favor use of *estar,* while Cortés-Torres (2004) determined that speakers with a lower level of education favor it. With respect to the factor of age in this community, Juárez-Cummings identified middle-aged (35–44) speakers as most strongly favoring *estar.* Meanwhile, in Puerto Rico, Brown and Cortés-Torres (2012) found that younger speakers (aged 20–29) favor *estar,* while Ortíz López (2000) determined that men favor it. Finally, in Costa Rican Spanish, Aguilar-Sánchez (2009) found that women and speakers with a lower level of education show a tendency to select *estar.* As these studies show, social factors play a crucial role in determining copula selection across different varieties of Spanish, although these differences are regionally specific. Due to this rich variation, copula selection merits further study in unexplored speech communities and among heritage learners.

3. Heritage Learner Characteristics

The increased attention paid to variable use and identity construction among heritage learners in the last two decades has already begun to provide linguists with a better understanding of language systems among early bilinguals with variable degrees of Spanish language ability. Particularly in the U.S. context, where heritage speakers' use of minority languages such as Spanish alongside English as well as learners' motivations, attitudes, and anxieties play a major role in defining how acquisition will occur and to what extent (Ducar 2012; Montrul 2008). From a more formal perspective, researchers have already begun to examine how heritage speakers' systems differ from those of native Spanish speakers, including a tendency to overuse masculine forms and produce non-native noun–adjective orderings (Cuza and Pérez-Tattam 2016). As a result of increased reliance on two contrasting linguistic systems, young learners in particular may eschew functional features from the L1 system in favor of a combined system drawing on both the L1 and L2. There is also no guarantee that heritage learners will become balanced bilinguals, with some researchers discussing the inevitable incomplete acquisition that heritage learners in contact situations may experience (Montrul 2008; Putnam and Sánchez 2013).[1] Whether due to reduced exposure in the home, reduced motivation to learn the heritage language, or an individual desire to instead show proficiency in the dominant language, there are a number of both linguistic and social factors that influence heritage language acquisition.

In addition, sociolinguistic researchers have begun to focus on factors such as language attitudes, social power, agency, and culture which may contribute to heritage language acquisition. Researchers have identified ways in which these findings apply to the language classroom. These factors serve as a means of indexing one's belonging to the social group of heritage language speakers. Ducar (2012) examined how attitudes toward the heritage language can motivate students to acquire the language that forms a part of their cultural background and how anxieties around speech are often reduced for heritage learners. She also emphasized the need for further quantitative research to better understand production tendencies. Lynch (2012) and Lynch and Avineri (2021) focused on topics such as the role of social power, agency, and culture in the heritage context to understand factors that propel heritage learners' acquisition. Díaz-Campos et al. discussed bilingualism in the context of heritage learning, focusing particularly on the function that linguistic features serve in conveying group identity.

3.1. Copulas and Heritage Learners

While the use of *ser* and *estar* has been the focus of research on native and L2 speakers, it has been studied less extensively in bilingual and heritage communities. Silva-Corvalán (1986, 1994, 2014), for example, examined the extension of *estar* in Mexican Spanish–English bilinguals with varying degrees of Spanish attrition. After studying all contexts in which *ser* and *estar* appeared in her data, she described the extension of *estar* as being accelerated by contact with English and its appearance in progressives. Silva-Corvalán and Montanari (2008) analyzed copula use by Spanish–English bilingual children and found that the use of *estar* does not reach levels comparable to those of L1 Spanish children until later in the acquisitional process, which they attributed to contact with English.

Furthermore, Aguilar-Sánchez (2009, 2012, 2017) found in his research on bilingual and multilingual speakers of English and Spanish that a higher level of contact with English is correlated with more frequent use of *estar*. He found that social variables, including age, level of education, gender, and level of bilingualism, predict copula selection. The higher the level of formal instruction in Spanish, the more the speakers avoid using *estar* with monolinguals. Ultimately, he proposed that contact with English accelerates linguistic change in bilingual communities.

A more recent study by Cuza et al. (2020) looked at distinctions in *ser* and *estar* and their development in English–Spanish bilinguals in the United States. They examined *ser*-favored and *estar*-favored adjectival predicates as well as contexts where *ser* is obligatory. Overall, they found that child heritage speakers produce a higher rate of *estar* in adjectival contexts that are typically associated with *ser,* while adult heritage speakers do not demonstrate such strong preference. However, in comparison with a group of monolingual Spanish speakers, both child and adult heritage speakers overextend their use of *estar*. Ultimately, the authors argued that this overextension is due to a lack of Spanish input and less use of the language compared to their monolingual counterparts.

These studies demonstrate how copula usage differs between heritage and monolingual Spanish speakers. More specifically, heritage learners—particularly children—tend to select *estar* in contexts in which *ser* is preferred by monolingual speakers. In addition, they demonstrate that social factors, especially the level of bilingualism and formal instruction in Spanish, influence copula choice. Ultimately, these studies indicate that there is considerable variation in the selection of *ser* and *estar* in the speech of heritage speakers.

3.2. Research Questions

Two questions guided our investigation into the use of Spanish copulas:

1. Do English L1 Spanish learners, non-English L1 Spanish learners, and heritage Spanish learners differ in their use of the copula verbs *ser* and *estar*?

Based on findings by Cuza et al. (2020), it was hypothesized that heritage learners would have the highest rate of *estar* usage among the groups. Meanwhile, it was proposed that English L1 learners of Spanish would favor *ser* as an analogue to the single English copula, while non-English L1 learners of Spanish were expected to show a greater mix of tendencies, depending on the norms of the variety of world languages they represented.

2. To what extent do the linguistic and extralinguistic factors identified for monolingual speakers of Spanish apply to the aforementioned groups of learners?

Much of the research surrounding copula acquisition and variationist research involves monolingual speakers of Spanish. While it is predicted that some of the linguistic factors will emerge as predictive of variable use among language learners and heritage speakers, particularly for the former group, it is expected that there will be a less complex selection system. English and non-English L1 learners of Spanish will likely focus on certain linguistic factors that they have been instructed in or begun to notice themselves, while heritage learners are expected to have a more native-like system of linguistic conditioning, which will influence their selection due to their increased contact with Spanish.

4. Methodology

Written data were drawn from the COWS-L2H corpus collected at the University of California, Davis from university language learning classrooms (Yamada et al. 2020). Participants, including both L2 learners and heritage speakers of Spanish, wrote class essays on several prompts, including descriptions of something "beautiful", "terrible", and "special", as well as memorable events such as favorite vacations. Composed of 1.3 million words, the corpus includes a longitudinal representation of 5057 short essays from 1846 student participants taking Spanish language courses at UC Davis (Davidson and Sagae 2022). As shown in Table 1, the corpus is composed of about one sixth heritage learner essays and one sixth essays by international students speaking English as an L2 and learning Spanish as an additional language, with the remaining two thirds of the corpus being made up of L1 English learners of Spanish.

Table 1. Breakdown of participants, essays, and words in the COWS-L2H corpus.

	Participants		Essays		Words	
Learner group	#	%	#	%	#	%
Heritage learners	346	18.3%	858	16.9%	242,000	18.8%
International learners with L2 English	361	19.1%	932	18.4%	218,000	16.9%
L1 English learners of L2 Spanish	1186	62.7%	3272	64.6%	829,000	64.3%
Total	1893		5062		1,289,000	

The essays in the corpus, which is digitized and publicly available online, were produced by learners at various levels of language proficiency, coming from introductory, intermediate, advanced, and heritage Spanish courses. The essays averaged 255 words and focused on several topics, including narratives and basic descriptions. While information was provided regarding learners' reading, writing, listening, and speaking proficiency, this was unfortunately not consistent within the corpus, and was omitted from the statistical analysis due to this lack of consistency.

Based on the data provided in the corpus, several additional contextual variables were included in the analysis. These included the aforementioned essay prompts as well as students' age, course level, language background, and experience studying abroad. A representation of international learners with L2 English is provided below in Table 2. Overall, as the objective of this study was to examine the use and patterns of acquisition and variation of *ser* and *estar,* these brief writings are ideal samples of language learners' output that provide ample tokens of both copulas.

Table 2. Breakdown of international learners' L1 by language.

Learner L1	#	%
Mandarin	172	47.6%
Vietnamese	43	11.9%
Arabic	15	4.2%
Hindi	15	4.2%
Cantonese	11	3.0%
Farsi	11	3.0%
Japanese	11	3.0%
Punjabi	11	3.0%
Gujarati	10	2.8%
Hmong	10	2.8%
Korean	10	2.8%
Other languages *	42	11.6%
Total	361	

* Languages with fewer than ten students.

All cases of the two copulas were extracted from a context preceding one of 127 adjectives that are considered to demonstrate variation in verb selection, taken from Díaz-Campos and Geeslin (2011). Following extraction, adjectives were excluded that did not show variable use with at least one case of both *ser* and *estar*. After determining the envelope of variation, which consisted of 8085 tokens, factors contributing to the selection of *ser* and *estar* were compared based on morphosyntactic, semantic, and social variables in mixed-effects logistic regression models. This modelling allows for a clear understanding of how heritage, English L1, and non-English L1 Spanish learners employ copulas and which linguistic factors mediate their use.

4.1. Dependent and Independent Variables

First, the dependent variable in the analysis was whether *ser* or *estar* was selected in each variable context. Adjectives were only included in the analysis if they were used with both copulas. Linguistic and extralinguistic factors were expected to predict variable selection among learners, although heritage learners were expected to show more complex patterns of variable use than Spanish language learners. L1 English speakers were expected to favor a single-copula system comparable to English, meaning that they would be more likely to favor *ser* overall, while heritage speakers have cultural and linguistic ties to Spanish that were predicted to influence selection.

Within the analysis, both linguistic and extralinguistic characteristics were studied in order to determine variable trends related to copula choice based on discussions found in previous studies. Four linguistic characteristics were selected: resultant state, adjective class, experience with referent, and frame of reference. Experience with referent was determined by examining the surrounding discourse context. Additionally, six social characteristics, including those provided by the corpus and those referenced explicitly in previous research on copula use, were included: age, gender, L1, experience with study abroad, prompt, and self-identification of proficiency. Table 3 provides a summary of the predictions for each of these variables based on trends identified in previous studies.

Table 3. Predictions for social and linguistic factors based on previous studies.

Factor	Direction of Effect	Investigation Providing Support
Adjective class	Mental and physical state → *estar* Observable traits → *ser* Status → *ser*	Silva-Corvalán (1986); Gutiérrez (1992, 1994); Díaz-Campos and Geeslin (2011); Juárez-Cummings (2014)
Experience with the referent	Indirect → *ser* Ongoing → *ser* Immediate → *estar*	Brown and Cortés-Torres (2012); Díaz-Campos and Geeslin (2011); Geeslin and Guijarro-Fuentes (2006, 2008)
Frame of reference	[- comparison] → *ser* [+ comparison] → *estar*	Brown and Cortés-Torres (2012); Díaz-Campos and Geeslin (2011); Gutiérrez (1994); Silva-Corvalán (1986)
Age	Puerto Rico 20–29 y/o → *estar* Mexico 35–44 y/o → *estar* Venezuela 46+ → *estar*	Brown and Cortés-Torres (2012); Díaz-Campos and Geeslin (2011); Juárez-Cummings (2014)
Gender	Costa Rica Women → *estar* Puerto Rico Men → *estar*	Aguilar-Sánchez (2009); Ortíz López (2000)
L1	Heritage speakers → *estar* L2 learners → *ser*	Silva-Corvalán (1986, 1994, 2014)
Study abroad experience	Heritage learners interact and benefit more if abroad. Language learners → more native-like	Potowski (2012); Pollock (2020)
Self-identification of proficiency	Heritage speakers at lower proficiency levels → comparable to intermediate/advanced L2 learners	Lynch (2012)

4.2. Data Analysis

Taking into consideration the predictions about these variables, the data were analyzed using a series of mixed-effects logistic regression models in the Rbrul software for R (version 4.0.2, Johnson 2009). This permitted a comparison of the influences conditioning

heritage and language learner selection of the copula, thereby determining the roles of linguistic and social variables in the use of *ser* and *estar* while controlling for adjective use through random effects. Based on the variable rule approach of sociolinguistics, this type of mixed-effects model reduces imbalance across naturalistic data caused by multiple speakers producing multiple tokens by using random effects (Johnson 2014).

These types of models provide three pieces of information that describe how variables influence selection, as discussed by Tagliamonte (2012). First, variables significantly predicting variation beyond the level of chance are indicated by a *p*-value of less than 0.05. Second, the direction of effect is shown by the ordering of the log-odds and weights for individual factors, representing the degree to which certain factors favor the application value; factor weights are from 0 to 1, with weights below 0.5 disfavoring the application value and those above 0.5 favoring it. Third, the effect magnitude is symbolized by the range of said factor weights, and shows the extent to which factor tendencies differ within individual social and linguistic variables. Model selection was carried out through a comparison of AIC and log-likelihood values, favoring simpler models with fewer variables where possible.

5. Results

This section presents the results of the mixed-effects logistic regression for the copula choice of L1 English learners of Spanish, L1 non-English learners of Spanish, and Spanish heritage speakers. As it was not initially clear whether these three groups would approach copula use similarly, three separate models were developed to compare linguistic and social variables within each group.

The first step in this process involved examining the linguistic and extralinguistic influences that predicted the use of *ser* and *estar* for L1 English learners of Spanish. From the corpus, 8085 tokens of the [copula + adjective] construction were obtained for this group of speakers; *estar* was used in 2058 constructions (i.e., 25.5%), while *ser* was used in 6027 (74.5%). In the mixed-effects model shown in Table 4, where the adjective form and speaker were held as random effects to account for possible individual and lexical variation, five variables were selected as conditioning copula selection for these learners. Use of *estar* was favored by a [+resultant] state, in individual comparisons, by students who had spent time abroad, in narrative essays, and by younger learners.

Table 4. Logistic regression of *estar* usage by L1 English learners of Spanish with the adjective and speaker ID treated as random effects.

Variable	Factor	Logodds	Tokens	% *Estar*	Factor Weight
Resultant State (*p* < 0.001)					
	+Resultant	1.521	228	81.1%	0.821
	−Resultant	−1.521	7857	23.8%	0.179
	range				*64.2*
Frame of Reference (*p* < 0.001)					
	Individual comparison	0.988	2560	53.1%	0.729
	Class Comparison	−0.988	5525	12.7%	0.271
	range				*45.8*
Study Abroad (*p* = 0.011)					
	Yes	0.261	542	29.2%	0.565
	No	−0.261	7543	25.2%	0.435
	range				*13*
Essay Prompt (*p* = 0.008)					
	Narrative	0.15	3188	34.8%	0.537
	Description	−0.15	4897	19.4%	0.463
	range				*7.4*
Age (*p* = 0.046)					
	continuous	logodds			
	+1	−0.039			

n = 8085, df = 8, Log-likelihood = −2371, AIC = 4760, R2Fixed = 0.117, R2Total = 0.678.

The next model examined non-English L1 learners of Spanish as broken down in Table 2; these were international students who identified English as their L2 and came from a variety of linguistic backgrounds. Among this group, 2166 tokens of the [copula + adjective] construction were identified, with 549 occurrences (25.4%) of *estar* and 1617 occurrences (74.6%) of *ser*. The mixed-effects model provided in Table 5 shows that three factors were predictive of copula use among these learners. For the resultant state, as with L1 English learners, [+resultant] favored *estar*; for frame of reference, again consistent with L1 English learners, individual comparisons favored this copula; and lastly, for adjective class, mental adjectives favored *estar* while status and observable traits disfavored it.

Table 5. Logistic regression of *estar* usage by non-English L1 learners of Spanish, with the adjective and speaker ID treated as random effects.

Variable	Factor	Logodds	Tokens	% *Estar*	Factor Weight
Resultant State ($p < 0.001$)					
	+Resultant	3	55	80.0%	0.953
	−Resultant	−3	2111	23.9%	0.047
	range				*90.6*
Frame of Reference ($p < 0.001$)					
	Individual comparison	1.166	621	59.6%	0.762
	Class Comparison	−1.166	1545	11.6%	0.238
	range				*52.4*
Adjective Class ($p = 0.027$)					
	Mental	1.43	1251	35.7%	0.807
	Status	−0.705	146	16.4%	0.331
	Observable Traits	−0.725	769	10.3%	0.326
	range				*48.1*

$n = 2166$, df = 7, Log-likelihood = −555, AIC = 1124, R2Fixed = 0.182, R2Total = 0.810.

Based on the congruity of non-heritage speakers, looking at both the raw percentages of *ser* and *estar* production and the similarities in resultant state and frame of reference, a third mixed-effects model was created that combined tokens produced by L2 learners of Spanish into a single cohort. In this grouping, there were a total of 10,251 [copula + adjective] constructions examined, with speakers using *estar* 25.4% of the time and *ser* 74.6% of the time. The resulting model in Table 6 incorporates significant variables from both previous models, yielding five in total. Use of *estar* was favored by a [+resultant] state, in individual comparisons, among mental and status adjectives, by students with experience studying abroad and, finally, in narrative essay prompts.

Table 6. Logistic regression of *estar* usage by all L2 Spanish Learners, with the adjective and speaker ID treated as random effects.

Variable	Factor	Logodds	Tokens	% *Estar*	Factor Weight
Resultant State ($p < 0.001$)					
	+Resultant	1.801	283	80.9%	0.858
	−Resultant	−1.801	9968	23.9%	0.142
	range				*71.6*
Frame of Reference ($p < 0.001$)					
	Individual comparison	1.003	3181	54.4%	0.732
	Class Comparison	−1.003	7070	12.4%	0.268
	range				*46.4*
Adjective Class ($p = 0.029$)					
	Mental	0.558	6033	33.2%	0.636
	Status	0.133	687	22.0%	0.533
	Observable Traits	−0.692	3531	12.8%	0.334
	range				*30.2*

Table 6. *Cont.*

Variable	Factor	Logodds	Tokens	% *Estar*	Factor Weight
Study Abroad (*p* = 0.004)					
	Yes	0.277	593	29.5%	0.569
	No	−0.277	9658	25.2%	0.431
	range				*13.8*
Essay Prompt (*p* = 0.004)					
	Narrative	0.143	4046	35.0%	0.536
	Description	−0.143	6205	19.2%	0.464
	range				*7.2*

n = 10251, df = 9, Log-likelihood = −2869, AIC = 5756, R2Fixed = 0.154, R2Total = 0.692.

Following the creation of this composite model, the Spanish heritage learners were examined to determine whether they too patterned with other learners. Despite producing more overall words than the international learners of Spanish, heritage learners had the lowest number of copulas produced in variable contexts, with 911 [copula + adjective] tokens. Of these, 258 constructions (28.3%) used *estar* and 653 (71.7%) used *ser*. In the mixed-effects logistic regression that was run on these tokens in Table 7, only two variables were found to predict copula selection. For frame of reference, individual comparisons were found to favor the use of *estar*; meanwhile, for course level, students in lower-level courses used higher rates of *estar*, although this tendency decreased according to ascending course level.

Table 7. Logistic regression of *estar* usage by Heritage Learners, with the adjective and student ID treated as random effects.

Variable	Factor	Logodds	Tokens	% *Estar*	Factor Weight
Frame of Reference (*p* < 0.001)					
	Individual comparison	0.694	357	42.0%	0.667
	Class Comparison	−0.694	554	19.5%	0.333
	range				*33.4*
Course Level (*p* = 0.025)					
	continuous	logodds			
	+1	−0.481			

n = 911, df = 5, Log-likelihood = −376, AIC = 762, R2Fixed = 0.060, R2Total = 0.650.

Understanding the Logistic Trends

To better understand the six variables shown to significantly predict copula selection between the composite group of Spanish learners and heritage learners, descriptive statistics were examined across the cohort. In this way, heritage learner language usage could be clearly contrasted to that of learners with a background in English, either as an L1 or as an earlier L2. The first variable described in all four models in the previous section was resultant state. As shown in Figure 1, heritage speakers had the lowest rate of *ser* with [−resultant] adjectives and the highest rate of *estar*. Meanwhile, their use of the two copulas in [+resultant] contexts reversed that of the Spanish learners, with much higher rates of *ser*. Meanwhile, the two groups of learners were nearly identical in their rates of copula production.

Next, for the frame of reference, Figure 2 demonstrates a less marked difference between the two Spanish learner groups as compared to heritage speakers. While learners had the highest rates of use of *ser* with class comparisons, heritage speaker rates were somewhat lower, while their use of *estar* was the highest. On the other hand, while learners used *estar* more for individual comparisons, heritage speakers showed considerably higher rates of *ser* use.

The following variable, adjective class, is illustrated in Figure 3. Here, heritage speakers can be seen to have the highest rate of *ser* use and the lowest rate of *estar* use for observable trait adjectives when compared to learners. Meanwhile, for status adjectives, heritage speakers have the highest rates of *estar*; for mental adjectives, usage is comparable across learner groups.
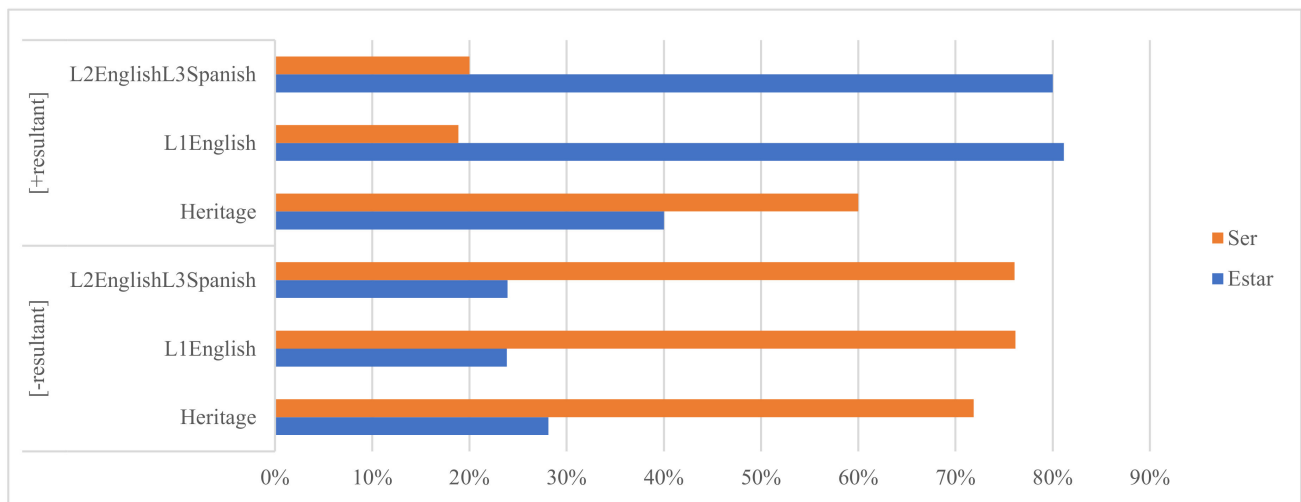
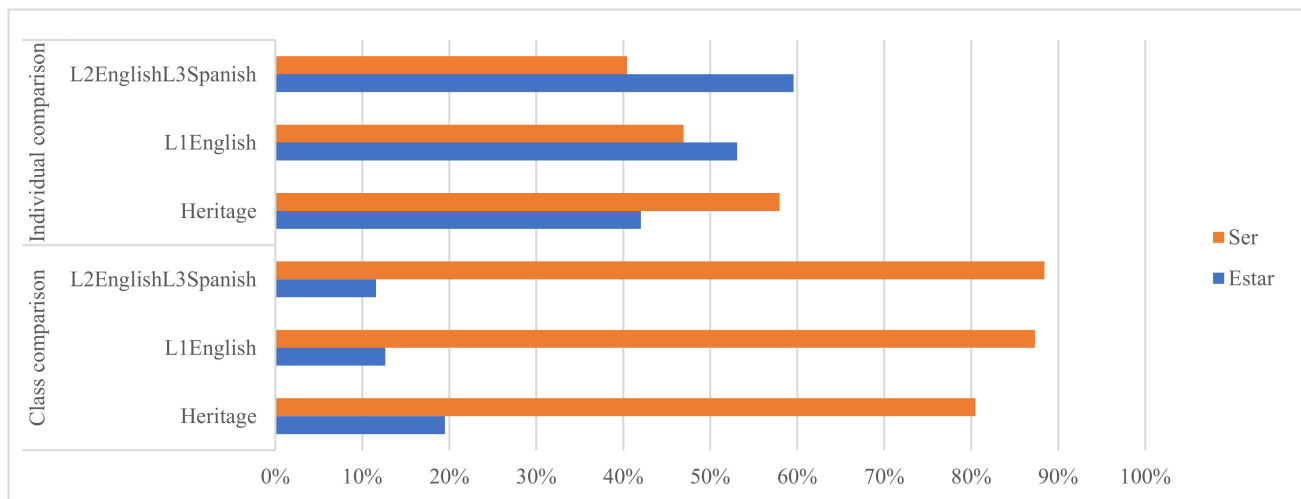**Figure 1.** Resultant state selection for each copula by speaker group.

Figure 2. Frame of reference selection for each copula by learner group.
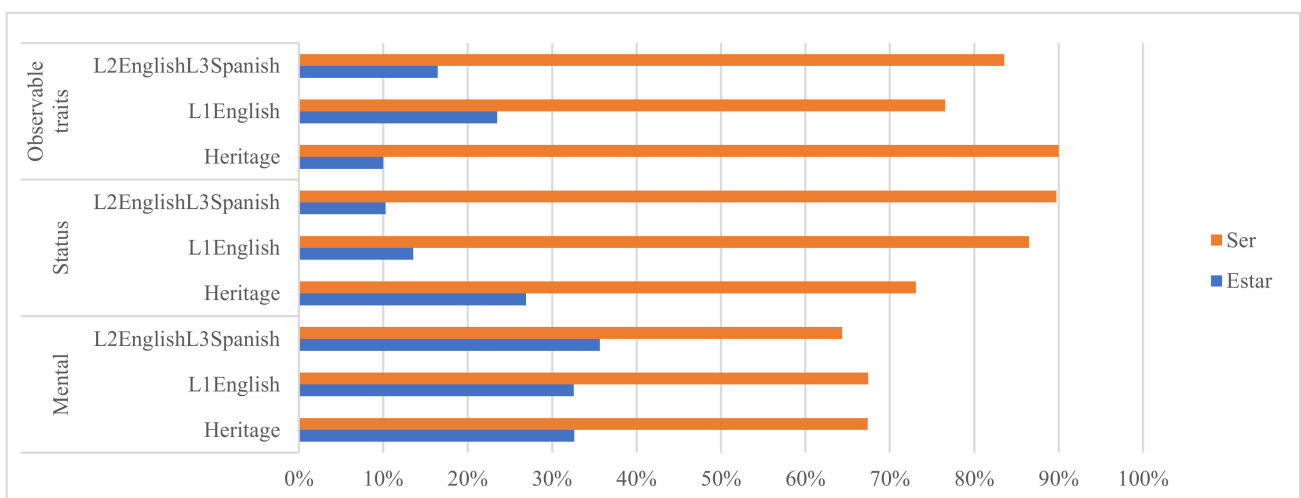

Figure 3. Adjective class for each copula by learner group.

The next significant variable examined by mixed-effects logistic regression was experience studying abroad (Figure 4). It can be noted that copula use among heritage learners

stays relatively stable regardless of their study experience abroad. However, among speakers who did not study abroad, *estar* use is at its highest among heritage speakers, whereas among those who did not, it is the L1 English and L1 non-English learners who use higher rates of *estar*.
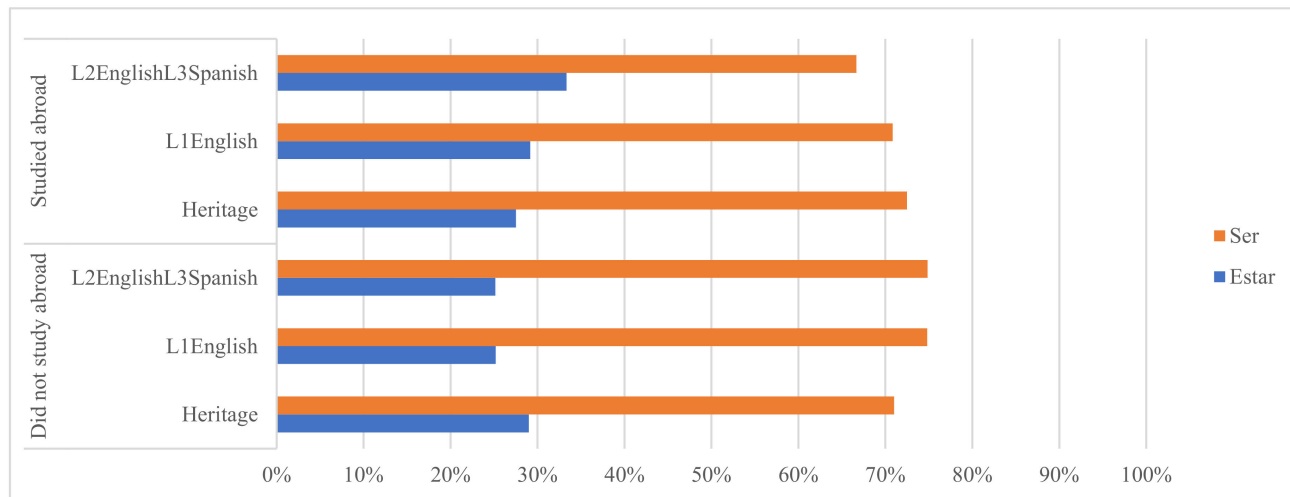
Figure 4. Study experience abroad for each copula by learner group.

The essay prompt variable was the next to predict copula selection, with the results shown in Figure 5. There is a particularly marked difference in use between learner groups in the description and narration tasks. While the former triggers extremely high rates of *ser*, the latter shows greater use of *estar*. Meanwhile, heritage learners demonstrate the same trend, although they show a considerably reduced difference by prompt type.
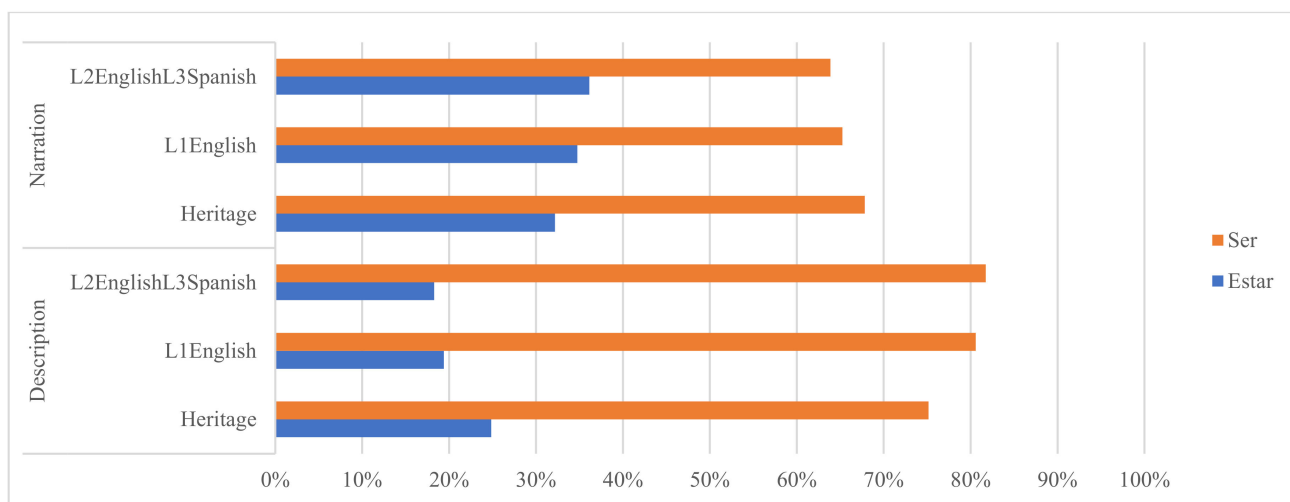
Figure 5. Essay prompt type for each copula by learner group.

Finally, the last significant variable selected in the mixed-effect logistic regression models as predicting copula selection was the course level of the speaker, where level one corresponds to the introductory Spanish courses at UC Davis and increases up to the most advanced fourth level (Figure 6). The trends here seem to suggest a type of overcorrection for non-heritage learners; while L1 English and international learners of Spanish have exceedingly low rates of *estar* use in essays taken from lower-level courses, those in the fourth level have an almost balanced system of copula use. Meanwhile, heritage learners start with some of the highest rates of *estar* usage, which drops in intermediate levels and reestablishes itself at about 35% among learners in the most advanced courses.
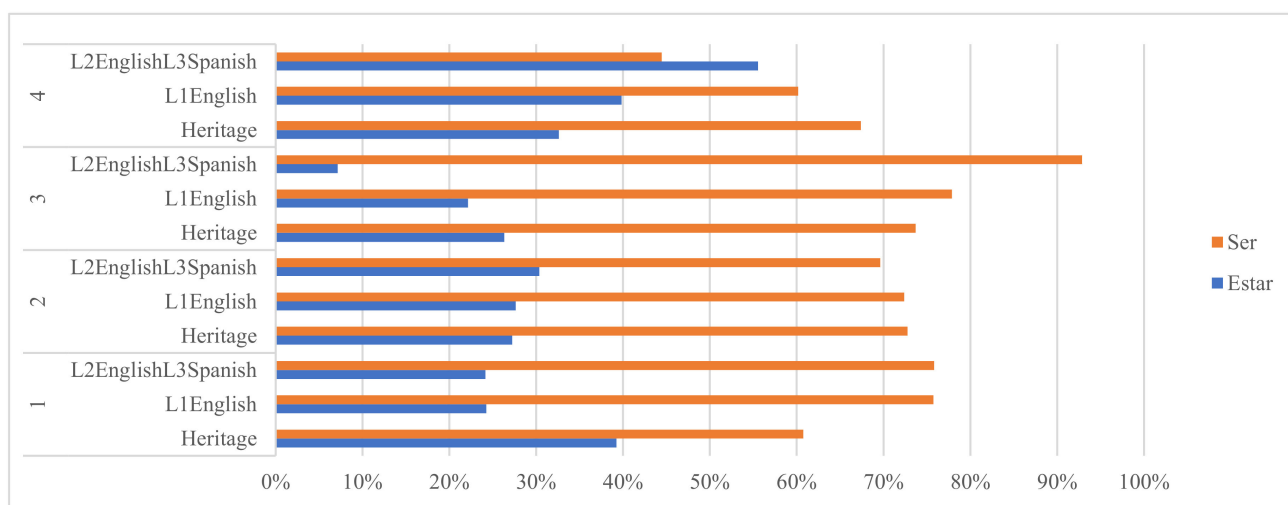
Figure 6. Course level for each copula by learner group.

6. Discussion

6.1. RQ1: Do English L1 Spanish Learners, Non-English L1 Spanish Learners, and Heritage Spanish Learners Differ in Their Use of the Copular verbs Ser and Estar?

Copula selection was found to be constrained by different factors in the analyses of English L1, non-English L1, and heritage Spanish learners. In samples produced by L1 learners of Spanish with English as their first language, both linguistic and extra-linguistic factors significantly predicted variation. From a linguistic perspective, a [+resultant] state and individual frame of reference favored the selection of *estar*. With respect to extralinguistic factors, *estar* was favored when participants had studied abroad, among younger participants, and in narratives. For learners of Spanish whose native language was not English, [+resultant] states and individual frames of reference favored use of *estar*. However, in this case, adjective class was shown to influence selection, with "mental" adjectives favoring *estar* and those categorized as "status" or "observable traits" disfavoring its appearance. In the mixed-effects logistic regression combining these groups, the extra-linguistic factor "course level" was selected as well, with a higher university course level coinciding with more use of *estar*, possibly indicating the role of instruction in copula selection.

Meanwhile, a different set of factors presented itself as influencing copula selection for heritage learners. Unlike with language learners, resultant state did not surface as a meaningful differentiator. Instead, only frame of reference and course level were selected, with individual comparisons and lower course levels favoring use of *estar*.

These results indicate that "frame of reference" is crucial to copula selection in all three groups, with *estar* being favored with individual comparisons; however, *estar* usage by L2 speakers is more frequent than that of heritage learners in cases in which there is an individual comparison. In an analysis of all L2 learners (Table 6), in cases in which an individual comparison is made, *estar* is selected in 54.4% of cases. Heritage learners, on the other hand, employ *estar* in 42% of these cases (Table 7). Based on an examination of factor weights and the descriptive results, although frame of reference is selected in every regression, L2 learners employ *estar* more, and this factor appears to play a larger role in copula selection.

While the resultant state was found to be significant for both L2 groups, it was not influential in copula selection among heritage learners. While there are great disparities in the selection of *estar* based on whether there is a resultant state in both L2 Spanish groups, as seen in Figure 1, there is a smaller difference between heritage speakers' rates of *estar* usage based on this.

In addition, copula choice is mediated by extra-linguistic factors in this dataset. For English-speaking L2 learners of Spanish, study abroad, age, and essay topic were selected as significant. Those students who had experience studying abroad and were younger

favored *estar,* suggesting that immersion and experience with native speakers contribute to *estar* use, as does a higher level of instruction. Additionally, heritage learners enrolled in higher-level university Spanish courses show an increment in the use of *estar,* providing evidence that exposure to Spanish in an academic setting promotes its use. This is consistent with the findings of Cuza et al. (2020), who noted that heritage learners have higher rates of *estar* in general when compared to L1 speakers of Spanish, as well as with Aguilar-Sánchez's findings that a higher level of formal Spanish instruction leads to heritage speakers adjusting their rates of *estar* usage.

6.2. RQ2: To What Extent Do Linguistic and Extralinguistic Factors Identified for Monolingual Speakers of Spanish Apply to Groups of Learners?

The current findings support the previous literature on L1 Spanish, showing that both linguistic and extra-linguistic factors mediate copula selection. First, the results for linguistic factors have been found in previous studies. Similar to the trends identified in the mixed-effects logistic regressions for both L2 groups, Geeslin (2003, 2005) and Díaz-Campos and Geeslin (2011) found that *estar* use is correlated with resultant state adjectives. Previous studies (Geeslin 2003, 2005; Brown and Cortés-Torres 2012; Gutiérrez 1992, 1994; Juárez-Cummings 2014; Ortíz López 2000) have found that *estar* often references mental and physical states, while *ser* describes observable traits. These findings mirror the results in the current study for L2 learners of Spanish whose native language is not English, with *estar* being favored to describe mental states.

The role of social factors in determining copula selection in native-speaking communities referenced in previous studies is less clear, with social stratification depending on the individual speech community. The results regarding course level seem to reflect the findings discussed in previous studies on speakers' level of education. Cortés-Torres (2004) found that speakers with a lower level of education favored the copula *estar* in a Mexican speech community. Similarly, Aguilar-Sánchez (2009) found that speakers with a lower level of education show a tendency to select *estar* in Costa Rican Spanish. The results in the current study regarding heritage learners show an association between lower course levels and increased use of *estar* among heritage learners.

The tendencies observed among Spanish heritage learners in the descriptive analysis can be linked to previous research on heritage learner motivations and identity construction. Considering that these speakers may well have a system that combines aspects of the L1 and the L2 (Putnam and Sánchez 2013), and that they have been found in the past to differ from native speakers (Cuza and Pérez-Tattam 2016), it makes sense that their production norms would be less unified than that of learners. While learners are receiving L2 input in Spanish for the first time and are given certain contexts and tools to aid in *ser* and *estar* use, heritage learners come into the language system with a pre-existing set of criteria for copula use. These criteria differ based on individual factors such as language exposure and proficiency and develop at different rates based on learners' attitudes toward Spanish learning and motivation to acquire the language (Ducar 2012; Montrul 2008). As such, copula use in this group is not a case of comparable decisions, as it is among English and international learners, instead consisting of variable individual approaches. This may help to explain why differences in copula use based on the resultant state are high for learners, who are taught to look out for this cue, while they remain low for heritage learners; similarly, for essay prompts, while English and non-English L1 learners may key on phrases that rely on *ser* or *estar,* heritage learners may be more likely to rely upon their own innate judgment. Thus, the reason behind the decrease in *estar* use may involve a more precise awareness of copula usage contexts.

The current study bolsters previous findings on L1 speakers, demonstrating that there is variation in copula selection in writing samples from L1 English, international, and heritage learners of Spanish. As in previous sociolinguistic studies, copula use is mitigated by a complex array of linguistic and extra-linguistic factors.

7. Conclusions

The present study investigated copula patterns found in the speech of heritage speakers in comparison to L2 learners with differing L1s. Based on the quantitative results, L2 learners seem to acquire patterns similar to those found in monolingual speakers, where pragmatic and semantic distinctions of the predicates determine the use of *ser* and *estar*. Meanwhile, the descriptive results paint a more complex picture in the case of heritage speakers, showing that these speakers have higher overall rates of *estar* use in variable contexts while those enrolled in higher course levels have lower rates of *estar* usage. Factors such as study experience abroad and the type of essay prompt did not have a marked effect on copula selection among heritage speakers.

This article provides an empirical analysis of corpora data of *ser* and *estar* in a population of language learners and heritage speakers, describing whether variable use of *ser* and *estar* follows the same trends as in monolingual varieties of Spanish described in previous research. The present comparative analysis allows us to understand heritage speakers' approach to variable copula choice. While L2 Spanish learners tend to follow trends identified in previous literature (e.g., Kanwit and Geeslin 2020), heritage speakers do not follow Caribbean and Central American norms (e.g., de Jonge 1993; Gutiérrez 1992; Díaz-Campos and Geeslin 2011).

As many of these speakers are second and third generation residents of the United States, their linguistic connections to other language contexts are reduced, with a much greater tendency to follow the norms identified among Latinx communities in the U.S. (Silva-Corvalán 1994) and to develop their own individual linguistic personas based on their unique social situation (e.g., Lynch and Avineri 2021). This is based on the type of input that these students are relying on; while heritage learners make use of community input, classroom learners are more likely to favor instructional input, causing the former to start off more closely resembling L1 norms. Future research could examine how individual differences and social factors in heritage speakers play a further role in copula selection.

Notes

[1] A similar phenomenon is referred to as Differential Acquisition by Kupisch and Jason (2016) and described by Fairclough (2005) as Second Dialect Acquisition (SDA).

References

Aguilar-Sánchez, Jorge. 2009. Syntactic Variation: The Case of Copula Choice in Limón, Costa Rica. Doctoral dissertation, Indiana University, Bloomington, Indiana.

Aguilar-Sánchez, Jorge. 2012. Formal Instruction and Language Contact in Language Variation: The Case of ser and estar + Adjective in the Spanishes of Limón, Costa Rica. In *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. Edited by Kimberly Geeslin and Manuel Díaz-Campos. Somerville: Cascadilla Proceedings Project, pp. 9–25.

Aguilar-Sánchez, Jorge. 2017. *Research Design Issues and Syntactic Variation: Spanish Copula Choice in Limón, Costa Rica*. Balti: Lambert Academic Publishing.

Batllori, Montserrat, and Francesc Roca. 2011. Grammaticalization of ser and estar in Romance. In *Grammatical Change: Origins, Nature, Outcomes*. Edited by Dianne Jonas, John Whitman and Andrew Garrett. Oxford: Oxford University Press.

Batllori, Montserrat, Francesc Roca, and Elena Castillo. 2009. Relation between Changes: The Location and Possessive Grammaticalization Path in Spanish. In *Diachronic Linguistics*. Edited by Joan Rafel Cufí. Girona: Universitat de Girona Publication Services, pp. 443–93.

Brown, Esther, and Mayra Cortés-Torres. 2012. Syntactic and Pragmatic Usage of the [estar + Adjective] Construction in Puerto Rican Spanish: ¡Está brutal! In *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. Edited by Kimberly Geeslin and Manuel Díaz-Campos. Somerville: Cascadilla Proceedings, pp. 61–74.

Camacho, José. 2012. 'Ser' and 'estar': Individual/stage level predicates or aspect? In *The Blackwell Handbook of Hispanic Linguistics*. Edited by José Ignacio Hualde, Antxon Olarrea and Erin O'Rourke. Oxford: Wiley-Blackwell, pp. 453–76.

Clements, Joseph Clancy. 1988. The semantics and pragmatics of the Spanish <COPULA + ADJECTIVE> construction. *Linguistics* 26: 779–822.

Clements, Joseph Clancy. 2006. Ser-estar in the Predicate Adjective Construction. In *Functional Approaches to Spanish Syntax*. Edited by Joseph Clancy Clements and Jiyoung Yoon. Basingstoke: Palgrave-McMillan, pp. 161–202.

Cortés-Torres, Mayra. 2004. ¿*Ser* o *estar*? La variación lingüística y social de estar más adjetivo en el español de Cuernavaca, México. *Hispania* 87: 788–95. [CrossRef]

Cuza, Alejandro, and Rocío Pérez-Tattam. 2016. Grammatical gender selection and phrasal word order in child heritage Spanish: A feature reassembly approach. *Bilingualism: Language and Cognition* 19: 50–68. [CrossRef]

Cuza, Alejandro, Nancy Rsaeyes, and Eduardo Lustres. 2020. Copulas *ser* and *estar* production in child and adult heritage speakers of Spanish. *Lingua* 249: 102978. [CrossRef]

Davidson, Sam, and Kenji Sagae. 2022. The UC Davis Corpus of Written Spanish, L2 and Heritage Speakers. GitHub. Available online: https://github.com/ucdaviscl/cowsl2h (accessed on 1 January 2023).

de Jonge, Bob. 1993. (Dis)continuity in language change: Ser and estar + age in Latin-American Spanish. In *Linguistics in the Netherlands [Linguistics in the Netherlands 10]*. Edited by Frank Drijkoningen and Kees Hengeveld. Amsterdam: John Benjamins, pp. 69–80.

Díaz-Campos, Manuel, and Kimberly Geeslin. 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. [CrossRef]

Díaz-Campos, Manuel, Molly Cole, and Matthew Pollock. in press. Sociolinguistic approaches to bilingual phonetics and phonology. In *The Cambridge Handbook of Bilingual Phonetics and Phonology*. Edited by Mark Amengual. Cambridge: Cambridge University Press.

Díaz-Campos, Manuel, Iraida Galarza, and Gibran Delgado. 2017. The sociolinguistic profile of ser and estar in Cuban Spanish: An analysis of oral speech. In *Cuban Spanish Dialectology: Variation, Contact, and Change*. Edited by Alejandro Cuza. Washington, DC: Georgetown University Press, pp. 135–60.

Ducar, Cynthia. 2012. SHL learners' attitudes and motivations: Reconciling opposing forces. In *Spanish as a heritage language in the United States*. Edited by Sara M. Beaudrie and Marta Fairclough. Georgetown: Georgetown University Press, pp. 161–78.

Fairclough, Marta. 2005. *Spanish and Heritage Language Education in the United States: Struggling with Hypotheticals*. Madrid/Frankfurt: Iberoamericana/Vervuert.

Geeslin, Kimberly L. 2003. A Comparison of Copula Choice in Advanced and Native Spanish. *Language Learning* 53: 703–64. [CrossRef]

Geeslin, Kimberly L. 2005. *Crossing Disciplinary Boundaries to Improve the Analysis of Second Language Data: A Study of Copula Choice with Adjectives in Spanish*. Munich: LINCOM Publishers.

Geeslin, Kimberly L., and Pedro Guijarro-Fuentes. 2006. The second language acquisition of variable structures in Spanish by Portuguese speakers. *Language Learning* 56: 53–107. [CrossRef]

Geeslin, Kimberly L., and Pedro Guijarro-Fuentes. 2008. Variation in contemporary Spanish: Linguistic predictors of *estar* in four cases of language contact. *Bilingualism: Language and Cognition* 11: 365–380. [CrossRef]

Gili Gaya, Samuel. 1961. *Nociones de Gramática Histórica Española*. Barcelona: Compendios de Divulgación Filológica.

Gutiérrez, Manuel J. 1992. The Extension of *Estar*: A linguistic change in progress in the Spanish of Morelia, Mexico. *Hispanic Linguistics* 5: 109–41.

Gutiérrez, Manuel J. 1994. Simplification, Transfer and Convergence in Chicano Spanish. *Bilingual Review/La Revista Bilingüe* 19: 111–21.

Hanssen, Federico. 1913. *Gramática Histórica de la Lengua Castellana*. Halle: Niemeyer.

Isasa, Ane Isasa. 2014. *Ser and Estar Variation in the Spanish of the Basque Country [Conference Presentation]*. Urbana-Champaign: Illinois Language and Linguistics Society 6.

Johnson, Daniel Ezra. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3: 359–83. [CrossRef]

Johnson, Daniel Ezra. 2014. Progress in Regression: Why Natural Language Data Calls For Mixed-Effects Models. Daniel Ezra Johnson. Available online: http://www.danielezrajohnson.com/ (accessed on 16 August 2023).

Juárez-Cummings, Elizabeth. 2014. Tendencias de uso de 'Ser' y 'Estar' en la Ciudad de México. *IULC Working Papers* 14: 120–37.

Kanwit, Matthew, and Kimberly L. Geeslin. 2020. Sociolinguistic competence and interpreting variable structures in a second language: A study of the copula contrast in native and second-language Spanish. *Studies in Second Language Acquisition* 42: 775–99. [CrossRef]

Kupisch, Tanja, and Rothman Jason. 2016. Interfaces with syntax in language acquisition. In *Manual of Grammatical Interfaces in Romance*. Edited by Susann Fischer and Christoph Gabriel. Berlin, Boston: De Gruyter, pp. 551–586.

Leonetti, Manuel. 1994. ser y estar: Estado de la cuestión. *Barataria* 1: 182–205.

Lynch, Andrew, and Netta Avineri. 2021. Sociolinguistic Approaches to Heritage Languages. In *The Cambridge Handbook of Heritage Languages and Linguistics*. Edited by Silvina Montrul and Maria Polinsky. Cambridge: CUP, pp. 423–48.

Lynch, Andrew. 2012. Key Concepts for Theorizing Spanish as a Heritage Language. In *Spanish as a Heritage Language in the United States*. Edited by Sara M. Beaudrie and Marta Fairclough. Georgetown: Georgetown University Press, pp. 79–97.

Marco, Cristina, and Rafael Marín. 2015. Origins and development of adjectival passives in Spanish A corpus study. In *New Perspectives on the Study of Ser and Estar*. Edited by Isabel Pérez-Jiménez, Manuel Leonetti and Silvia Gumiel-Molina. Amsterdam: John Benjamins, pp. 239–66.

Montrul, Silvina. 2008. *Incomplete Acquisition in Bilingualism. Re-Examining the Age Factor*. Amsterdam: John Benjamins.

Montrul, Silvina. 2016. *The Acquisition of Heritage Languages*. Cambridge: Cambridge University Press.

Ortíz López, Luis A. 2000. Extensión de *estar* en contextos de *ser* en el español de Puerto Rico: ¿evolución interna o contacto de lenguas? *Boletín de la Academia Puertorriqueña de la Lengua Española*, 98–118.

Pollock, Matthew. 2020. Did you say peso or beso?: The perception of prevoicing by L2 Spanish learners. In *Variation and Evolution: Aspects of Language Contact and Contrast Across the Spanish-Speaking World [IHLL 29]*. Edited by Sandro Sessarego, Juan J. Colomina-Almiñana and Adrián Rodríguez-Riccelli. Amsterdam: John Benjamins Publishing, pp. 127–61.

Potowski, Kim. 2012. Identity and Heritage Learners: Moving Beyond Essentializations. In *Spanish as a heritage language in the United States*. Edited by Sara M. Beaudrie and Marta Fairclough. Georgetown: Georgetown University Press, pp. 179–99.

Putnam, Michael T., and Liliana Sánchez. 2013. What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism* 3: 478–508. [CrossRef]

Silva-Corvalán, Carmen, and Simona Montanari. 2008. The acquisition of *ser*, *estar* (and be) by a Spanish-English bilingual child: The early stages. *Bilingualism: Language and Cognition* 11: 341–60. [CrossRef]

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

Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. [Oxford studies in language contact]. Oxford: Clarendon.

Silva-Corvalán, Carmen. 2014. *Bilingual Language Acquisition: Spanish and English in the First Six Years*. Cambridge: Cambridge University Press.

Tagliamonte, Sali A. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Hoboken: Blackwell Publishing.

Valdés, Guadalupe. 2001. Heritage languages students: Profiles and possibilities. In *Heritage Languages in America: Preserving a National Resource*. Edited by Joy Kreeft Peyton, Donald A. Ranard and Scott Mcginnis. Washington, DC: Center for Applied Linguistics/Delta Systems, pp. 37–77.

Yamada, Aaron, Sam Davidson, Paloma Fernández-Mira, Agustina Carando, Kenji Sagae, and Claudia Sánchez-Gutiérrez. 2020. COWS-L2H: A corpus of Spanish learner writing. *Research in Corpus Linguistics* 8: 17–32. [CrossRef]