The use of Spanish Ser and Estar + adjectives: A sociolinguistic pilot study on the oral Spanish of Costa Rica

Abstract

The study of copula choice in Spanish in contact situations (i.e., among Spanish-English bilinguals) in contexts where Spanish is a minority language (i.e., Southwestern United States) has suggested that the copula use is undergoing change. Part of this change is that the copula estar is becoming more accepted in an extended pre-adjectival context especially with adjectives of size, physical appearance, age, and those of description and evaluation and this change has been attributed to contact with English (Silva-Corvalán, 1986, 1994). This claim has been challenged by Gutiérrez (1992, 1994) because he found that monolingual Mexicans from the same social class as those studied by Silva-Corvalán (1986, 1994) showed the same behavior with regard to copula choice. Diaz-Campos & Geeslin (2004) found similar results regarding copula choice in an analysis of the spoken Spanish of Caracas, a language with no contact with English.

Furthermore, the study of the Spanish copula variation is becoming a prolific field in both first language (Díaz-Campos & Geeslin, 2005; Silva-Corvalán, 1986, 1994; Gutiérrez, 1992, 1994) and second language studies (Geeslin, 2000, 2001, 2002). However, studies of the Spanish copula have been done mostly in settings where Spanish is the only language spoken (i.e., monolingual Spanish) and are described as monolingual societies (e.g., Caracas) and/or Spanish is a minority language (i.e., Los Angeles), and in the acquisition of Spanish as a foreign or second language in similar settings (i.e., Texas).

One setting that has yet to be explored is that of largely monolingual Spanish communities which are in contact, directly or indirectly, with minority languages (Thomason, 2001) and are located within small geographical boundaries. Costa Rica is an example of such a context (Aguilar-Sánchez, 2005a; 2005b). Because of the differences in the claims resulting from copula choice studies in both grammatical (i.e., copula + adjective) and geographical (i.e., contact with other languages) contexts, exploring copula choice in a context where Spanish is a majority language and English is a minority language will help us to better understand the processes of copula choice. The type of language contact found in such a small geographical area such as Costa Rica helped us achieve this goal.

The present work studies oral interviews to ten highly educated Costa Ricans living in part of Costa Rica traditionally described as monolingual. It was found that variables such as adjectival class, gradience, predicate type, experience with the referent, and resultant state are strong predictors of copula choice. These findings are consistent with previous findings such as those by Geeslin and Diaz-Campos (2005). True variation settings were found using statistical methods and analysis used to study variation in language. Finally, evidence in support of the idea that copula choice is a stable change (Díaz-Campos and Geeslin, 2005) was also found among the group of speakers under study.

I would like to thank Dr. Manuel Díaz-Campos, Dr. Kimberly Geeslin for their insightful comments on the reviewing of this paper. All interpretations are my own. Dan McNerly and Paul Kanczuzewski for his support and editing. And all the people who helped me make this paper a reality.
Introduction

The study of copula choice in Spanish in contact situations (i.e., among Spanish-English bilinguals) in contexts where Spanish is a minority language (i.e., Southwestern United States) has suggested that copula use is undergoing change. Part of this change is that the copula *estar* is becoming more accepted in an extended pre-adjectival context especially with adjectives of *size, physical appearance, age,* and those of *description and evaluation*. This change has been attributed to contact with English (Silva-Corvalán, 1986, 1994). This claim has been challenged by Gutiérrez (1992, 1994) because he found that monolingual Mexicans, in Mexico, from the same social class and geographical region as those studied by Silva-Corvalán (1986) showed the same behavior with regard to copula choice. Díaz-Campos & Geeslin (2004) found similar results regarding copula choice in an analysis of the spoken Spanish of Caracas, a language with no direct contact with English.

This area of research has become a prolific field in both first language (Díaz-Campos & Geeslin, 2004, 2005; Silva-Corvalán, 1986) and second language studies (Geeslin, 2000, 2001, 2002). However, these studies have been done mostly in settings where Spanish is the only language spoken (i.e., monolingual Spanish) and are considered to be monolingual societies (e.g., Caracas); and/or settings where Spanish is a minority language (e.g., Los Angeles), and in the acquisition of Spanish as a foreign or second language in similar minority language settings (e.g., Texas).

Even when it has been categorized as a monolingual society at large, Costa Rica is not different from most multilingual societies in the world (Aguilar-Sánchez, 2005a; 2005b). This claim is supported by Thomason’s (2001) claim that “… no nation is monolingual in the sense that all its citizens speak one and only one language, and probably no nation is monolingual in the less trivial sense that everyone who lives there speaks the same language natively” (36); and by works in bilingualism such as Romaine (2004, 1996). Therefore, Costa Rica can be added to the list of countries that Thomason (2001) uses to support her claim. These countries have only a single official language, but they also have minority languages within their boundaries (e.g., Australia, Brazil, Cambodia, among others). Due to the fact that Costa Rica’s constitutionally official language is Spanish, most of the linguistic phenomena induced by language contact have not yet been paid attention to. That being the case, it is important to point out that the language situation in Costa Rica, with regard to cases such as this one, is more complex than what is normally believed. Thus, the linguistic situation of Costa Rica cannot be categorized as that of a monolingual society. Like most of the countries in Central America, Costa Rica is a country with native English speakers in a small percentage (about 2%) of its population (Aguilar-Sánchez, 2005a, 2005b) who are mostly inhabitants of the province of Limón.

Due to both methodological reasons and the belief that the situation in Limón is an isolated phenomenon particular to a small region in Costa Rica most studies have been done, within the field of Creoles and on the description of linguistic and/or cultural phenomena (Herzfeld, 2002, 1994, 1992, 1991, 1983, 1980a, 1980c, 1978a, 1978b, 1977, Melendez Chaverri & Duncan, 1974; Winkler, 1998; Winkler & Obeng, 2000). However, the influence, or lack thereof, of Limonesque English on the Spanish of Costa Rica and specifically on the Spanish of the province of Limón has not yet been investigated. This lack of literature on the impact of language contact in the area of Limón and/or of Costa Rica, due to its linguistic diversity, in general constitutes a primary reason for undertaking this and the subsequent studies in this area.

In this work, I present previous work done in the study of Spanish regarding copula choice in the *copula + adjective* construction, and the results of an analysis of oral data from 10
Costa Ricans from the *Corpus de Habla Culta*. The present work focuses on the examination of oral data from this particular corpus to fill in the gap caused by the lack of a description of the copula + adjective construction of Costa Rican Spanish. The results of this study will serve as a comparison point for a bigger study that follows it. Furthermore, the results are accompanied by a discussion of the implications of the findings and the effect that these have on the description of copula + adjective use. This will serve as a point of comparison to shed light on the concepts of change and language contact, the categorization of the uses of *ser* and *estar*, and the relationship with degrees of language contact found.

My focus in the next sections is to introduce a description of the previous research with regard to copula choice in Spanish. But before I carry this task out, it is useful to provide the reader with a concise background of the overall linguistic situation found in Costa Rica.

**Costa Rica: A multilingual society**

A profile of the languages spoken in Costa Rica is displayed here to present the reader with a short, but comprehensive, account of the linguistic situation in Costa Rica.

Costa Rica’s other languages are believed to be undergoing changes caused by contact with standard Spanish. Many of these native languages are near extinction, or are already extinct. Within its small territory, we find, in Costa Rica, seven Amerindian languages (6 alive and 1 extinct), six transplanted languages, a sign language, and the official language (see Map 1). According to the Ethnologue report from the Summer Institute of Linguistics (SIL) (2005) the Amerindian languages include languages from the Chibchan/Talamancan, the Chibchan/Rama, and the Oto-Manguean/Chiapanec-Mangue families. Within the Chibchan/Talamancan family we find Boruca spoken by about 35 people in a population of about 1000 Borucas; Bribri spoken by about 11,000 people out of 12,172 Bribris; Cabécar spoken by about 7,072 in a population of 9,308 Cabécarens; and Teribe spoken by about 5 people in a population of approximately 300 Teribes. From the Chibchan/Rama family we find Maléku Jaika spoken by about 750 speakers out of a population of 1,074 Malékus. Furthermore, from the Oto-Manguean/Chiapanec-Mangue we find Chorotega which unfortunately has been extinct for many years although there is a population of about 795 Chorotega Indians from the Guanacaste region near the Nicaraguan border that still carry out their traditions and keep some lexical items from Chorotega.

**Map 1**

**Linguistic Map of Costa Rica**

Source: www.Ethnologue.com

Furthermore, Costa Rica’s linguistic profile is enhanced by the presence of six transplanted non-Amerindian languages. These languages are found in Costa Rica as a result of work-related diaspora, more specifically with the construction of the railroad to the Atlantic by

Work-related diaspora brought English to Limón. Nowadays, it is spoken by about 55,100 Afro-Costa Ricans mainly in the province of Limón. The diaspora also introduced two Chinese dialects Ngābēre and Yue Chinese. These varieties are spoken, in the entire country, by approximately 5,092 and 4,500 people respectively. In addition, with the contemporary immigration came Plautdietsch from the Germanic language family spoken by about 100 people in the northern area of the country; Basque and Eastern Yiddish (no speaker population numbers found).

As we can see, Spanish in Costa Rica has been in contact with other languages since the early years of colonization. In this paper, I try to describe the Spanish copula choice in copula + adjective constructions by highly educated Costa Ricans with ages ranging from 28 to 50 years old. This study will provide a point of departure for a future investigation focusing on the contact situation in the Limón region.

Two sections related to the topic follow to provide the reader with a short but comprehensive account of what has been done in the study of copula choice in Spanish. The first section is a summary of theoretical accounts regarding the definition, function, and proposed theoretical approaches that had lead to the description and structure of the variables used to study the copula + adjective construction. The second section contains previous work done regarding copula choice in the fields of sociolinguistics and second language acquisition.

Theoretical accounts on copula + adjective

Early studies on copula distinction date back to the late XIX and early XX centuries. During that time ser was defined as predicates of permanent conditions and estar as predicates of non-permanent conditions. One of the earliest of the works that argue against this view is that of Andrade (1919) in which he states that the course of evolution of ser and estar consists in the restriction of the original meaning of the former with regard to the extension of the latter and that we must look for the present elements of distinction in the latest developments of estar. He defines estar as being associated with the characteristic feelings which attend immediate perceptions and their representations and ser as related to concepts of judgments. He further states that estar is related to “warmth” and “intimacy” which no object of mere conception ever attains and that ser is related to the “colder” logical relations of the objects (22). He also states that in judging affective elements our point of view must always be the purpose or the emotional state of the speaker/writer.

Morely (1925) set forward the notion that no one underlying principle suffices to explain all the uses of ser and estar. However, he states that the contrast is usually between durative and transitory, or between act and state. Morely (1925) also states that emphasis could be laid on one point of view or another; and that in some uses two ideas come together. Morely (1925) argues that language is “a creature of feeling and not of reason” and that since the XVII century and before, estar has invaded the field once held by ser and that there is reason to believe that the invasion has not entirely stopped. Andrade’s (1919) and Morely’s (1925) claims are the early stages for the formation of the variable experience with the referent in which the speaker could have an immediate (i.e., momentary impression), ongoing (i.e., previous knowledge), and/or indirect (i.e., second-source knowledge) experience with the referent. This variable is of a
discursive nature and it is one of the first discourse-related variables used in the study of copula choice.

Bolinger (1944) follows the discussion by supporting Bull’s (1942) principles for some Spanish equivalents of “to be” stating that objectivity must be thrown out as a final test for _ser_ and _estar_. He suggests that objective permanence influences the average human being to form subjective concepts that will call for the use of _ser_. That objectivity may itself be something that needs to be verbalized for which we would use _ser_. Bull (1942) defends the thesis that the nature of the attributes as means of determining whether _ser_ or _estar_ would be used before a predicate adjective should be taken into account in the analysis of copula choice. Thus, limiting adjectives and those negative ones that cannot be expressed by the positive of another adjective are used only with _ser_ (e.g., _es limitante “Be-3/E-Sig limiting”) and that the other group of adjectives may be used either with _ser_ and _estar_ (es/está inteligente “BE-3/E-Sig intelligent”). Bull’s (1942) claim has given rise to the notion of _adjective class_ as a possible predictor of copula choice in Spanish. _Adjective class_ refers to whether the adjective is an adjective of _age, size, and physical appearance_, among others (i.e., the semantic nature of the adjective).

Bolinger (1944) also proposes, as an alternative to Bull’s (1942) model, that the nature of the subject is an important feature on the choice between _ser_ and _estar_. Bolinger (1944) puts forward three types of subjects: infinite, infinitesimal, and evolutionary. The former is composed of metaphysical entities such as _Dios “God”, amor “love”, fortuna “fortune”, santidad “sanctity, “_ and other concepts when taken abstractly. This type of subject generally calls for the use of _ser_; however, they call for _estar_ only when they refer to everyday experiences. The second category is composed of events of very limited duration such as vuelo “flight”, _verbal nouns_, and _noun clauses_. _Ser_ is used to indicate a first impression or a normal average concept (e.g., _fue un vuelo corto “it was a short flight”_). In this case we would be saying that the first impression is the only impression we get because of the limited duration of the event. The last category is composed of facts of daily experience. They have duration and a life history. These entities call for the use of _ser_ or _estar_ depending on whether the attribute is regarded as normal (e.g., _el niño es muy bueno “the child is well behaved”_ or as a change from the norm (e.g., _el niño está muy bueno hoy “the child is well behaved today”_). From these accounts the notion of the nature of the _subject_ as a variable for the study of copula choice is derived and has yielded different approaches to the classification of it that differ from the ones proposed by Bolinger (1944) and will be explained in a subsequent section.

The following claims have been the basis for the construction of the variable _frame of reference_ which establishes whether there is comparison between the referent and him/herself at a different period of time or between him/herself and a group.

Regarding this topic, Moellering (1944) states that “the meaning of an adjective is a resultant of reciprocal influence between itself and the noun with which it is linked” (598). Moellering (1944) does this in an attempt to enhance the utility of the principle of subjective interpretation as a basis of grammatical analysis. Moellering (1944) also stresses the notion that an adjective’s role in a noun-adjective construction is not that of modifying the noun but of singling out some aspect of the actual object to which reference is being made and focusing attention on that aspect. Thus, when choosing between _ser_ and _estar_, when the attribute is predicated, the speaker indicates with _ser_ that this attribute is one of a permanent nature and with _estar_ the speaker indicates that this attribute is phase-like rather than permanent to that fixed set of related attributes which are his present concept of the subject.
De Mello (1979) puts forward the idea that a clear understanding of the distinction between *ser* and *estar* is best achieved by considering them according to their functions, as principal verbs, auxiliary verbs or attributive (copulative) verbs. I will focus on the latter function. De Mello (1979) argues that the distinction between *ser* and *estar* as attributive verbs can be found in the semantics of each verb. De Mello’s (1979) thesis is that *ser* has no semantic value in its attributive role and *estar* has a value beyond that of its function as an attributor. Thus, the difference between *ser* and *estar* in their attributive functions lays on the difference between [-semantic value] for *ser* and [+semantic value] for *estar*. This distinction may explain why certain adjectives change meaning with the use of one or the other copula. Furthermore, De Mello (1979) introduces the notion of *change* when he states that *estar*, in contrast to the semantically empty *ser*, indicates that the idea of change is present in the mind of the speaker. Change is the manifestation of something more basic in the mind of the native speaker of Spanish. This manifestation is basically the concept of temporal duration or the passage of time.

Similar to the notion of change set forth by DeMello (1979), Franco and Steinmetz (1983, 1986) introduce the notion of comparison. The comparison could be said to be between the individual and a group of individuals or between the individual and him/herself at a different point in time. They affirm that the selection of *ser* and *estar* + adjective could be based in subjective factors and in the perception of the external phenomenological world. They emphasize that it is wrong to state that only *estar* indicates a comparison or contrast; and they put forward the thesis that either copula can indicate a comparison and/or contrast. Thus, *Ser* is related more with a comparison between the referent and a group whereas *estar* is related to a comparison between the referent and him/herself at a different point in time.

In addition to his contribution to the notion of frame of reference, De Mello (1979) introduces the notion of resultant state. This notion distinguishes the use of the copula + past participle as passive voice when used with *ser* and that of a resultant state when used with *estar*.

What follows is a summary of the discussion which has given rise to the variable predicate type. Predicate type refers to whether the interpretation of the predicate is limited in time or not. First, we find the **Stage-Level/Individual-level hypothesis** (Kratzer, 1995; Carlson, 1977). This hypothesis states that the copula *ser* combines with the individual-level predicates (ILPs). Individual-level predicates express (more or less) permanent or essential properties such as inteligente “intelligent”. The hypothesis also states that *estar* combines with stage-level predicates (SLPs) which express temporary or accidental properties such as cansada “tired-FEM”. Maierborn (2005) argues against the **Stage-level/Individual-level hypothesis**. Her argument is based on the notion that *ser* and *estar* both display the same lexical semantic properties, *estar* differing from *ser* only in presupposing a relation to a specific discourse situation. Meierborn (2005) found reasons to reject a view of the stage-level/individual-level distinction as a grammatical phenomenon with a conceptual foundation. She states that what is at the heart of the *ser/estar* alternation is a specificity presupposition on the topic situation. Thus, *estar* is the discourse-dependent variant of *ser*. She states that this discourse dependency is lexically triggered by *estar*, structurally resolved by means of the functional category of aspect, and finally, pragmatically licensed through some kind of topic situation contrast. Meierborn (2005) claims that by using *estar* a speaker restricts his or her claim to a specific discourse situation, whereas by using *ser* the speaker makes no such restriction.

Similarly, Clements (2006) states that instead of being predicate types as defined by the Stage-level/Individual-level hypothesis, there are two readings that the copula use gives to the
copula + adjective construction. The first one is an Individual Level reading which refers to a defining characteristic of the referent. The second is Stage Level reading which is a comparison of the characteristic of the referent not against others but against itself at an earlier time.

With this refined claim in mind, Clements (2006) introduces a new variable to the study of copula choice. This variable is related to the internal making of the adjective. Clements (2006) states that there are three types of adjectives: one without underlying dynamic situation (e.g., mortal “mortal”) and two types with underlying dynamic situation: event adjectives and process adjectives. The latter two types are further subdivided as unidirectional (e.g. bajo “short” → alto “tall” [unidirectional process]; soltero “single” → casado “married” [unidirectional event]) and bidirectional (e.g. sincero “sincere” ←→ insincero “insincere” [bidirectional process]; levantado “out of bed” ←→ acostado “in bed” [bidirectional event]). Ser is used exclusively with adjectives with no underlying dynamic situation. Both ser and estar are used with adjectives with underlying bidirectional processes, unidirectional processes, bidirectional events, and unidirectional events; but the reading given to the predicate changes according to the copula used.

Clements (2006) proposes that this factor helps us to predict the choice of copula in Spanish as follows: adjectives with no underlying dynamic situation will appear exclusively with ser, will have an individual-level reading, and are said to be ill-formed with estar (e.g., El humano es/esta* mortal “The human being is mortal”). Bidirectional-process adjectives may appear with ser with an individual-level reading (e.g., el niño es gordo “the child is fat”) and with estar with a stage-level reading (e.g., el niño está gordo “the child is fat”). Unidirectional-process, bidirectional-process, and bidirectional-event adjectives will have an individual-level reading with ser (e.g., el tren está viejo “the train is old”) and stage-level reading with estar (e.g., el tren está viejo “the train is old”). Unidirectional-event adjectives will have an stage-level reading with estar (e.g. el hombre está casado “the man is married”) and an individual-level reading with ser (e.g., el hombre es casado “the man is married”) only if the referent is animate. Finally, bidirectional-event adjectives have a stage-level reading with estar (e.g., el niño está levantado “the child is out of bed”), but are said to be ill formed as an attributive construction (e.g., el niño es levantado “the child is out of bed”). This construction is only grammatical as a passive voice construction for which the meaning would be “the child is picked up by someone”, a distinction put forward by DeMello (1979) as described above.

Another variable that may play a role in the prediction of copula choice that has not been included in models is gradiencty. Gradiencty refers to the classification of adjectives in two groups. One group is that of non-gradient adjectives or absolute adjectives (i.e., adjectives that cannot be found in “more or less” constructions such as casado “married”) and the other is that of gradient adjectives (i.e., adjectives that can be found in “more or less” constructions such as mojado “wet”). Givón (1984) affirms that adjectives come in antonymic pairs. These pairs can either be in a gradual-degree relation to one another (e.g., seco “dry” – mojado “wet”) or in an absolute relation to one another (e.g., vivo “alive” – muerto “dead”). The adjective type seco “dry”– mojado “wet” corresponds to activity and stative verbs, in Vendler’s (1967) classification of verbs. The type vivo “alive” – muerto “dead” corresponds to non-graduating states which can be said to be analogous to the notion of achievements. Absolute adjectives will pair more with ser because of the permanent nature and the telic context they create. For this group of variables the prediction is that gradient adjectives, because they are less absolute states, will pair more with ser; but if they will pair with estar, they create a durative context. These predictions
are similar to the results, found in previous studies, of the variable resultant state (i.e., adjectives derived from verbal forms such as *mojado* "wet").

In sum, from Andrade’s (1919) and Morely’s (1925) claims we derived the variable experience with the referent in which the speaker could have an immediate (i.e., momentary impression), ongoing (i.e., previous knowledge), and/or indirect (i.e., second-source knowledge) experience with the referent. From Bolinger (1944) we derived the nature of the subject as a possible predictor of copula choice. Bull’s (1942) and Bolinger’s (1944) claims include the notion of adjective class as a possible predictor of copula choice. As mentioned above, adjective class refers to the nature of the adjective whether it is an adjective of age, size, or physical appearance, among others. The variable frame of reference which establishes whether there is comparison between the referent and him/herself at a different period of time or between him/herself and a group is derived from Bolinger (1944), Moellering (1944), DeMello (1979), and Franco and Steinmetz (1983, 1986). Moreover, DeMello (1944) also introduces the variable resultant state which distinguishes the use of the copula + past participle as passive voice when used with *ser* and that of a resultant state when used with *estar*. From Kratzer, (1995), Carlson (1997), Maierborn (2005) and Clements (2006) we derived the variable predicate type/reading which refers to whether the interpretation of the predicate is limited in time or not. Finally, from Givón (1984) we obtain the variable gradience which refers to the classification of adjectives in two groups. One group is that of non-gradient adjectives or absolute adjectives (i.e., adjectives that cannot be found in "more or less" constructions such as *casado* "married") and the other is that of gradient adjectives (i.e., adjectives that can be found in "more or less" constructions such as *mojado* "wet").

Although theoretical accounts regarding copula choice in Spanish have provided a type of spinal cord to the study of copula choice, further studies are needed in order to model what seems to be a structure with multiple faces. Two fields that have seen a blossoming of models that explain the use of copula choice is that of second language acquisition of Spanish and that of Sociolinguistics. Both fields have used the variables described above in search of empirical evidence for the prediction of copula choice in Spanish.

In the following section, the reader is presented with a summary of the discussion regarding copula choice in Spanish within the fields of sociolinguistics and second language acquisition. This is done to contextualize the scope of the present study within the bigger picture of sociolinguistics.

**Previous Work on the variation of the structure *ser* and *estar* + adjective**

In the field of sociolinguistics and second language acquisition several studies on the copula choice in Spanish have shed light on the variation that exists in speakers’ choice in this context. However, this variation is not new to grammarians of either the American tradition (i.e., American scholars) or the Spanish tradition (i.e., Spanish or Latin-American scholars) as it was presented in the previous section. Due to the fact that the discussion on copula choice has revolved around the notion of a change in progress, I will focus, in this section, on presenting the reader with a review of the relevant studies in this area. Three subsections follow. The first one deals with monolingual Spanish, the second one with studies done in second language acquisition of Spanish, and the third one with studies in settings where Spanish is in contact with English.
Monolingual Spanish

Both sociolinguistics and second language acquisition have been able to inform each other, as proposed by Geeslin (2002a, 2002b), and this can be seen from the next set of studies of copula choice. An example of this type of work is Geeslin and Díaz-Campos (2005) in which the authors apply the model proposed by Geeslin (2000b) to study the copula choice in a variety of Spanish said to have limited or no contact with English. The authors studied the Spanish of Caracas, Venezuela. They used the Estudio sociolingüístico de Caracas (1987) [financed by the Consejo de Desarrollo Científico y Humanístico de la Universidad Central de Venezuela] which includes half-hour interviews, conducted in 1987 and 1988 with 160 speakers born and raised in Caracas, with parents also from Caracas. Each interview was divided evenly among four age groups, both genders and five socioeconomic groups. The authors studied data from four age groups: 14-29, 30-45, 46-60, and above, three social classes (low, middle and upper); and they were evenly divided between men and women (2 participants per cell). The variables predicate type, susceptibility to change, experience with the referent, resultant state, adjective class, copulas allowed, age, and social class were predictors of copula choice. They also found that older speakers tended to favor the use of estar, a pattern of behavior that is an indication of the stability of this phenomenon in Caracas Spanish.

Díaz-Campos and Geeslin (2005) studied whether the extension of estar was a change in progress or a stable change. They studied how copula use differed between older and younger speakers by studying the frequency of use of estar and the predictors of it. They also studied how frequency of occurrence of the predictors of estar would vary across age groups. They studied the speech of 48 speakers from the Estudio Sociolingüístico de Caracas (1987) from four age groups and three social classes. Their sample was evenly distributed between men and women.

The authors found, by using separate cross-tabulations of each linguistic predictor of estar across time and $\chi^2$, that older generations tend to favor the use of estar and there are no prominent differences between female and male speakers. In their predictor analysis by age group they found that predicate type, resultant state, adjective class and copulas allowed were common predictors for all age groups. However, there were different predictors that were significant for specific groups. These predictors were susceptibility to change for age group 14-29; experience with the referent for groups 14-29, 46-60, and 61+; and socioeconomic class for group 46-60. In this study, gender was never selected as statistically significant. They concluded that a sociolinguistic interview does not elicit the same types of contexts from all speakers, and that three of the variables show a greater proportion of the categories that favor estar for older speakers.

Their results revealed that discourse factors such as predicate type, resultant state, adjective class, and copulas allowed are more important predictors than age and gender.

Díaz-Campos and Geeslin (2005) studied the extension of estar with adjectives and its relationship to language contact. They examined the individual categories of the adjective class variable to determine their relationship with the use of estar. Their intention was to relate adjective classes to the innovative use found in previous work. They stated that the classes that are the best predictors of use of estar may not actually represent innovation. They were guided by questions such as what categories of the adjective class variable showed the most frequent use of estar; what categories of the adjective class variable show innovation; and how their results relate to previous studies on adjectives and copula choice. In a new analysis of the data they
found that adjective classes with highest use of estar were mental state, size, physical appearance, color, status, and age. The distribution of estar with each adjective class across the variable copulas allowed showed more tokens of estar in the both allowed category than were required. Their results revealed uses of estar in contexts formerly reserved for ser for all age groups. A high frequency of use of estar is especially noticed with adjectives of age, size, and status. Furthermore, when both copulas are allowed speakers tended to favor estar, evidence that innovative uses are emerging. They concluded that the monolingual Venezuelan Spanish shows similarities to US/Mexican Spanish in the innovative use of estar with adjectives of age, size, physical description, evaluation, color for certain age group; for adjectives of status and size for all age groups; and age and description of personality for the older speakers.

Second Language Acquisition of Spanish

Following the tradition of studying copula choice and guided by the need to find a model that explains the features that triggered variation among copula choice in second language acquisition, Geeslin (1999) pioneered the study of copula choice with regards to the copula+adjective environment in the field of second language acquisition. Because the nature of the present work is purely sociolinguistic, I will focus on the elements relevant to the study of copula choice and those elements related to native speakers presented in Geeslin’s (1999) and subsequent works.

Geeslin (2000) proposes an innovative approach to copula choice. She proposes a study of usage based on second language Spanish data. The author asked which linguistic contextual features are significant predictors of the usage of estar at all levels of development and which linguistic contextual features serve to describe change across time as learners progress towards the acquisition of copula choice. She found that adjective semantic features such as class, semantic transparency, animacy, susceptibility to change, dependence on experience among others are strong predictors of copula choice among second language speakers of Spanish.

Furthermore, Geeslin (2002a) tried to unite two fields of linguistic investigation: sociolinguistics and second language acquisition. In order to allow both fields to inform each other, the author proposes two hypotheses. The first one is that the examination of individual categories of the semantic transparency variable will show that copula choice varies from one category to another and that each category of semantic transparency will predict copula choice at all levels. The second hypothesis predicts that categories within semantic transparency will provide a new way to address learner accuracy and avoid difficulties faced in previous research. Seventy-seven high schoolers with limited exposure to native Spanish participated in this study. Data were collected via a background questionnaire, a guided interview, a picture description task, and a contextualized questionnaire. The author found that indeed copula choice varies from one category to another and that each category of semantic transparency predicts copula choice at all levels. Her results show that dynamicty, a variable related to aspect, was a significant predictor of estar at all levels. Grammatical accuracy and task type were also significant predictors at all levels of proficiency. Even though these variables were significant at all levels of proficiency, there were some other variables that were significant at different levels of proficiency. These variables included directionality (i.e., whether the adjective was unidirectional such as age, bidirectional such as emotional state, or non-directional such as race) was significant at lower levels. Similarly, susceptibility to change was a predictor at the lowest level of proficiency whereas frame of reference was a significant predictor at the highest levels of proficiency with modality and telicity. The author concludes that each of the variables
included in the model requires individual analysis to further specify their role in learner grammar because each variable was significant at least once, therefore, providing new research questions and directions for future study.

Geeslin (2002b) tried to compare the paths of language change with those of language acquisition. The author states the hypothesis that the stage set forth for language change may mirror stages of SLA exhibited in her study such that constraints on susceptibility to change will be acquired prior to the frame of reference constraint. The results showed that the path of acquisition of copula choice mirror those of language change because the latter is moving toward simplification (Silva-Corvalán, 1991; Myhill, 1991) and the former is moving away from simplification to a more native-like grammar (Geeslin 2001). This means that learners move from a one-to-one mapping of copula meaning to a more complex system whereas native speakers are moving from a complex system of copula use to one where both copulas are being found in contexts of linguistic variation.

Geeslin (2003) studied copula choice among ten native speakers of Spanish from different countries. The author did this in order to explore the issue of coding for accuracy in studies of copula choice by learners of Spanish as a second language. Geeslin (2003) found that responses on a questionnaire instrument, designed to identify those contexts in which only ser is allowed, where only estar is allowed, and contexts in which both copulas are allowed, were not homogeneous or categorical. In fact, more than half of the items showed variation and demonstrated that the issue of addressing accuracy is not limited to a small sample of contexts. The author concludes that because of the variation found in the answers from native speakers of Spanish, a different type of analysis is needed. Instead of the traditional error analysis, an analysis of features might offer more insight to understand copula choice.

Spanish in contact with English
Silva-Corvalán (1986, 1994) explored the extension of estar in the speech of 27 bilinguals of different generations and degrees of Spanish language attrition. The data were collected through recordings of conversations between the researcher and the participants. There were three groups of participants. Those who were born in Mexico and migrated to the U.S.A. after the age of eleven represent Group I; those who were born in the U.S.A. or migrated to the U.S.A. before the age of six represent Group II; and those born in the U.S.A. whose parents were of Mexican ancestry who fit the definition of the second group represent Group III.

Silva-Corvalán (1986) claims that the innovation found in her study represents part of an evolutionary trend in Spanish and other Romance languages and that language contact accelerates this trend. She states that the condition of reduced access or lack of access to formal varieties of the language must be met in order for this diffusion to happen and that internally motivated changes which involve generalizations across languages are also accelerated in a situation of extended contact. With regard to copula choice, Silva-Corvalán (1986) suggests that the extension of estar in progressives (i.e., estar + present participle) and its frequent association with be (i.e., be+v-ing) in those constructions may favor the rapid diffusion of estar in the context of predicate adjectives, where Spanish has evidenced a slow process of change independent of any language-specific influence (604). She suggests that there is no noticeable movement toward a steady functional specialization; but that the continuous renovation of Spanish due to the arrival of other immigrants keeps Los Angeles Spanish from changing entirely. She concludes by saying that the Spanish copula opposition with attributes is lost to a large extent among speakers in her Groups II and III, but that these speakers are unlikely to pass on Spanish to their descendants. This led her to assume that the system described in her study
which maintains a limited amount of meaningful variation will persist as a defining feature of English-dominant bilinguals. This claim has been challenged by Gutiérrez (1992, 1994) because he found that monolingual Mexicans from the same social class as those studied by Silva-Corvalán (1986, 1994) showed the same behavior with regard to copula choice. Gutiérrez (1992, 1994) and Silva-Corvalán’s (1986, 1994) findings are evidence that further studies in the area of language contact with regard to copula choice are needed.

Summary

The review of relevant previous literature regarding the copula+adjective construction points to four important issues. First, in order to better understand copula choice in Spanish; we need to use an approach that takes into account different linguistic and social features such that put forward by Geeslin (2000b) in her study of second language acquisition of Spanish; and Geeslin and Díaz-Campos (2005) and Díaz-Campos and Geeslin (2004, 2005) in sociolinguistics. Second, there is still debate in the field with regard to the variables that help us to predict copula choice in monolingual Spanish (Díaz-Campos and Geeslin, 2004, 2005). However, research and theoretical accounts have shown that the following variables are strong predictors of copula choice in Spanish: experience with the referent in which the speaker could have an immediate (i.e., momentary impression), ongoing (i.e., previous knowledge), and/or indirect (i.e., second-source knowledge) experience with the referent; adjective class which refers to the nature of the adjective whether it is an adjective of age, size, and physical appearance; among others; frame of reference which establishes whether there is comparison between him/herself at a different period of time or between him/herself and a group; resultant state which distinguishes the use of the copula + past participle as passive voice when used with ser and that of a resultant state when used with estar; predicate type/reading which refers to whether the interpretation of the predicate is limited in time or not; and gradience which refers to the classification of adjectives in two groups non-gradient adjectives or absolute adjectives (i.e., adjectives that cannot be found in “more or less” constructions such as casado “married”) and gradient adjectives (i.e., adjectives that can be found in “more or less” constructions such as mojado “wet”). Third, that monolingual and contact Spanish show similarities with regard to copula choice (Silva-Corvalán, 1986, 1994; Gutiérrez, 1992, 1994; Geeslin and Díaz-Campos, 2005) show that more studies are needed. Finally, that there is a lack of evidence related to copula choice in contexts in which Spanish is the majority language and English is the minority language. All four issues serve as the basis for the undertaking of the present work. Furthermore, the evidence presented here points towards the belief that the phenomenon under study is a stable change; and our goal, then, should be to discover what processes underlie the categorization of copula choice by native speakers in this specific context: copula + adjective.
The present study

Previous studies regarding copula + adjective construction have studied the influence of English in Spanish, the development of the copula system, and a stable change in varieties with no contact. The present work studies the phenomenon of copula choice, but in a setting that has been as yet unexplored: Spanish from Costa Rica. As mentioned in the review of the literature, Costa Rican Spanish has been in contact with other languages since the early years of the colonization. Thus, this study investigates the differences in copula choice by highly educated Costa Ricans. Thus, in search of an appropriate model to the study of copula + adjective construction in Spanish, I followed Geeslin’s (2000a, 2000b), Geeslin’s and Díaz-Campos’ (2005) and Díaz-Campos’ and Geeslin’s (2004, 2005) overall suggestions to the study of copula choice.

From the previous review of the literature and description of the present study, three research questions are set forward:
1. What linguistic and social variables help predict the use of ser and estar?
2. Are the variables the same for the prediction of ser as the ones for estar or are they different?
3. Is Spanish from highly educated Costa Ricans different from those of speakers of the same age from other varieties of Spanish? More specifically, is there evidence of differences in copula choice?

Methodology

Participants

The data pool consists of 10 oral interviews from the “Corpus de Habla Culta” of Costa Rica. These speakers are university educated Costa Ricans who hold professional positions in the country and represent the cultured class of the country. They are evenly distributed between men and women and their ages range from 28 to 50 years old.

Procedure

The data come from half hour Labovian interviews conducted in 1987. These interviews included topics such as childhood memories, near-death experiences, and the present political situation of the country.

Data coding

For the data coding, Geeslin’s (2003) and Díaz-Campos’ and Geeslin’s (2005) coding schema is used in order to assure comparability between studies and the possible generalization of the findings.

These variables include the dependent variable which is the use of ser or estar in the copula + adjective construction. The independent variables include predicate reading (Geeslin and Díaz-Campos’ predicate type) which refers to whether the interpretation is limited in time and the reading given to the predicate is individual-level (e.g., ‘Hay unos momentos que son individuales’ [There are moments that are individual/personal] [m34]) or stage-level (e.g., [los profesores] son los mejores’ [the teachers are the best] [h30]). Susceptibility to change which refers to whether the quality of the referent could change (e.g., entiendo que el decreto es tajante [I understand that the law is strong], h30). Frame of reference which refers to whether there is a comparison of the referent implied (e.g., ‘un par de potreros que eran especiales’ [A pair of pasture fields that were special] [b46]; ‘pero soy muy desconfiada’ [but I am very distrustful] [m37]). Experience with the referent which answers the question of whether the speaker has first-hand knowledge (e.g., fue tan curioso llegar a saber [it was very curious to get to know] [h30]) and whether it is an on-going or immediate reaction (e.g., yo soy católico [I am catholic]...
Resulant state which answers to the question whether the adjective is a resultant state (e.g., *esa estructura estaba construida* [that structure was built] [h37] or not. Adjective class which refers to what semantic class that best describes the adjective (e.g., *porque habría sido muy simpático haber inventado* [because it would have been very nice to have invented] [h37]. I will not include the variable copulas allowed because, as Geeslin (2003) points out regarding the coding of accuracy, “a researcher who codes for accuracy would also be required to make judgments based on his or her own variety of Spanish, without taking into account the varied input learners receive” (718). Even though this paper does not study learner’s data, this restriction applies to native speaker data because coding for accuracy (i.e., whether one or both copulas are allowed) would be to impose the researcher judgment on the data being studied. Even though the original purpose of this variable was to study the environments in which both copulas can or are used regardless of the prescriptive rules and to determine which adjectives are found with both copulas within studies of second language acquisition of Spanish where the environments were controlled for (Geeslin 2000, 2001, 2003), in the present study the same cannot be said because the researcher is not able to control for the specific environments where copulas can appear. Thus, a search for places where true variation is found is sought by statistical methods. Two logistics regressions, one with *ser* and one with *estar* as the application value, are used to determine if the same factors that predict the appearance of *ser* would predict the appearance of *estar* or if they are different factors and where these factors differ.

Following Clements (2006) theoretical approach to the study of copula+adjective construction, three new linguistic independent variables were included. The first one is subject which distinguishes between first-order-entities (i.e., people, animals, plants, and things) (e.g., *como M. es tan santa* [because M. is such a saint] [m46]) and second-order-entities (i.e., states, processes, and events, more likely to be located in time and to take place rather than exist) (e.g., *pero es bonito (el trabajar)* [but working is nice] [m49]). This variable is included because it has not been tested before. Geeslin (2002) tested the variable animacy with regards to the subject of the construction and found it not to be a significant predictor for copula choice. The variable subject in this study analyzes the subject not in terms of animacy, but in terms of type of entity it represents for what Clements (2006) propose the following predictions: first order entities are more commonly found with both copulas whereas second order entities are more commonly found with *ser*.

The other two variables are semantic characteristics of the adjective. As we recall form the review of the literature Clements (2006) states that there are three types of adjectives: one without underlying dynamic situation (e.g., *mortal “mortal”*) and two types with underlying dynamic situation: event adjectives and process adjectives. The latter two types are further subdivided as unidirectional (e.g. *bajo “short”* → *alto “tall”* [unidirectional process]; *soltero “single”* → *casado “married”* [unidirectional event]) and bidirectional (e.g. *sincero “sincere”* ← *insincero “insincere”* [bidirectional process]; *levantado “out of bed”* ← *acostado “in bed”* [bidirectional event]). His theoretical analysis of this category predicts that adjectives with no underlying dynamic situation and with an individual-level reading appear exclusively with *ser* and are said to be ill-formed with *estar*. Bidirectional-process-adjectives may appear with *ser* and have an individual-level reading and with *estar* and have a stage-level reading. Unidirectional-process-adjectives and bidirectional-process & bidirectional-event-adjectives will have an individual-level reading with *ser* and a stage-level reading with *estar*. Unidirectional-event-adjectives will have an stage-level reading with *estar* and, if the referent is animate, it will have an individual-level reading with *ser*. Finally, bidirectional-event-adjectives
have a stage-level reading with estar and are said to be ill-formed with ser. Therefore, it is important to test these predictions empirically.

The last variable is gradience which refers to whether the adjective is gradient or not. In other words, whether the adjective is non-gradient or absolute which means that they cannot be found in “more or less” constructions such as casado “married” or whether they are gradient adjectives which can be found in “more or less” constructions such as mojado “wet”. Givón (1984) predicts that gradient adjectives, because they are less absolute states, will pair more with ser and with estar creating a durative context. On the other hand absolute adjectives will pair more with ser because of the permanent nature and the telic context they create. These predictions are similar to the results, found in previous studies, regarding the variable resultant state. In the present study the variable gradience is entered in the model in order to test its predictive effect against the variable resultant state.

The last of the independent variables included in the present study were of a non-linguistic nature. This variable was gender. Gender refers to the biological gender of the participant: male or female. Gender is the only non-linguistic variable included here because one of the purposes of this paper is to study an age bracket where variation is said to exist the most (i.e., mid-age group). Because it is a corpus of “Habla Culta”, only the middle class with higher education is represented. Table 1 presents a summary of the linguistic variables to be entered in the model for the present study.

### Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Categories</th>
<th>Example</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicate reading</td>
<td>[- Stage-level]</td>
<td>Hay, Elena está enferma ‘Today, Elena is sick’</td>
<td>Is the interpretation limited in time?</td>
</tr>
<tr>
<td>Susceptible to change</td>
<td>[- Susceptible]</td>
<td>Elena es simpática ‘Elena is nice’</td>
<td>Can the quality of the referent change?</td>
</tr>
<tr>
<td>Frame of reference experience with the referent</td>
<td>[- Comparison]</td>
<td>El niño está alto ‘The boy is (grew) tall’</td>
<td>Is a comparison of the referent implied?</td>
</tr>
<tr>
<td>Resultant state</td>
<td>[- Resultant]</td>
<td>La situación está mala ‘The situation is bad’</td>
<td>Does the speaker have first-hand knowledge?</td>
</tr>
<tr>
<td>Adjective class</td>
<td>[- Size]</td>
<td>Joven ‘young’</td>
<td>Which semantic class best describes the adjective (in the sense it is used in the given context)? [note: both ‘description’ categories are last resorts]</td>
</tr>
<tr>
<td>Subject</td>
<td>[- first order]</td>
<td>Juan es joven ‘John is young’</td>
<td>Does the subject referent exist?</td>
</tr>
<tr>
<td>Underlying structure</td>
<td>[- Bidirectional Process]</td>
<td>Juan está mojado</td>
<td>Does the adjective have an underlying dynamic situation? What type?</td>
</tr>
<tr>
<td>Gradiency</td>
<td>[- gradient]</td>
<td>El hombre es mortal</td>
<td>Can the adjective be seen in a “more or less” gradient?</td>
</tr>
</tbody>
</table>

1 Part of this table was taken from Geeslin and Díaz-Campos (2005).
Data Analysis

The data were processed utilizing a VARBRUL analysis through GoldVarb X (Sankoff, Tagliamonte, and Smith 2005) for Windows. This is a binomial regression analysis of the factors included in the study. The results show factors that predict the source of errors encountered in the corpus with their corresponding log-likelihood and significant level as well as their probability weights. This type of analysis has been used in studies within the variationist sociolinguistic framework (Díaz-Campos, 2002; 2003, 2004; among others). The information yielded by this type of analysis has helped the variationist sociolinguistic field move forward towards more precise accounts of language change.

Results and Discussion

Results

In this section I present the results yielded by the VARBRUL analysis for both ser and estar. First, the results of the regression and the probabilistic weights by variable and their importance are presented. Second, a discussion of the results is presented.

A total of 337 tokens were collected and coded following the criteria in Table 1 [page 29]. From these, 263 (78%) were tokens of ser and 74 (22%) were tokens of estar. All nine variables were introduced into the regression analysis. From these nine variables, the exact same predictive groups were selected as significant predictors whether ser or estar were selected as application values. The results are presented with estar as the application value since we want to determine the factors favoring innovation. Variables eliminated in the analyses include frame of reference, susceptibility to change, subject, gender, and underlying event. Table 2 summarizes the results. Table 2 is organized as follows: the left column shows the factor groups that were

<table>
<thead>
<tr>
<th>Factor Group</th>
<th>Factor</th>
<th>n</th>
<th>%</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 2: Adjective class</td>
<td>Description/Evaluation</td>
<td>14/197</td>
<td>7</td>
<td>0.315</td>
</tr>
<tr>
<td></td>
<td>Status</td>
<td>3/47</td>
<td>68</td>
<td>0.771</td>
</tr>
<tr>
<td></td>
<td>Mental/physical state</td>
<td>16/25</td>
<td>64</td>
<td>0.931</td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td>5/17</td>
<td>29</td>
<td>0.736</td>
</tr>
<tr>
<td></td>
<td>Description of a person(ality)</td>
<td>1/33</td>
<td>3</td>
<td>0.411</td>
</tr>
<tr>
<td></td>
<td>Physical appearance</td>
<td>5/17</td>
<td>29</td>
<td>0.820</td>
</tr>
<tr>
<td>Group 4: Gradiency</td>
<td>Gradient</td>
<td>38/212</td>
<td>17</td>
<td>0.398</td>
</tr>
<tr>
<td></td>
<td>Non-gradient</td>
<td>36/125</td>
<td>28</td>
<td>0.668</td>
</tr>
<tr>
<td>Group 5: Predicate reading</td>
<td>Stage-level</td>
<td>38/59</td>
<td>64</td>
<td>0.909</td>
</tr>
<tr>
<td></td>
<td>Individual</td>
<td>36/278</td>
<td>12</td>
<td>0.380</td>
</tr>
<tr>
<td>Group 8: Experience with the referent</td>
<td>Ongoing</td>
<td>39/219</td>
<td>17</td>
<td>0.463</td>
</tr>
<tr>
<td></td>
<td>Indirect</td>
<td>25/107</td>
<td>23</td>
<td>0.469</td>
</tr>
<tr>
<td></td>
<td>Immediate</td>
<td>10/11</td>
<td>90</td>
<td>0.985</td>
</tr>
<tr>
<td>Group 9: Resultant state</td>
<td>Non-resultant</td>
<td>31/274</td>
<td>11</td>
<td>0.389</td>
</tr>
<tr>
<td></td>
<td>Resultant</td>
<td>43/63</td>
<td>68</td>
<td>0.877</td>
</tr>
</tbody>
</table>

*Log Likelyhood = -80.602 p = .014

The variable Adjective class was selected as the first predictor. Figure 1 shows that adjectives of size, physical appearance, mental/physical state, and status favor the use of estar (e.g., no es porque estén estéticamente lindas ‘it is not because they are aesthetically pretty’ [m30]) whereas adjectives of description/evaluation and description of a person(ality) disfavor it (e.g., el sentir de la mujer es importante ‘women’s feelings are important’ [m34]).
The second variable group selected as a significant predictor was gradiency. Figure 2 shows that, in term of gradiency, gradient adjectives disfavor the use of estar (e.g., *Esa palabra es muy problemática* ‘that word is very problematic’ [h28]) while non-gradient adjectives favor it (e.g., *pero si están bien estas tres* ‘but these three are pretty’ [m30]). These results are in contradiction to Givón’s (1984) predictions. We will go back to this issue in the discussion section where we explore the relationship between gradiency and other factors.

The third variable selected as significant predictor of the use of *estar* was predicate reading. Figure 4 shows that a stage-level-reading of the predicate favors the use of *estar* (e.g., [*los adolescentes] están sueltos* ‘adolescents are free’ [m46]) whereas an individual-level reading of the predicate disfavors it (e.g., *que es muy bonita* ‘that is very pretty’ [m49]). These results agree with those of previous studies (Geeslin 2000, 2001, 2003) and theoretical accounts (Clements 2006, Meierborn 2005) with regard to predicate type (reading) in both first and second language data.

The fourth factor group selected as a significant variable was experience with the referent. Figure 5 shows that ongoing and indirect experience with the referent disfavor the use of *estar* (e.g., *en el sentido de saber lo que es bueno* ‘in the sense of knowing what is good’ [when describing the personality of youth/people] [m34]) whereas immediate experience with the referent favors it (e.g., *otros estaban pálidos* ‘others were pale’ [when narrating about an accident he was in] [h37]). The results for this factor group also agree with the results of previous studies in copula-adjective choice.
The fifth variable group selected as significant was resultant state. Figure 6 shows that non-resultant states disfavor the use of estar (e.g., como profesor fui muy jovial ‘as a teacher I was very joyful’ [h46]) whereas resultant-states favor it (e.g., que uno esta medio dormido ‘that one is half asleep’ [h37]). Again, it is worth mentioning here, that DeMello (1944) made a distinction in the use of ser with a resultant state as a passive voice; therefore, this particular issue imposes a restriction on the use of ser with resultant states which may explain why resultant states favor the use of estar and not of ser. This distinction also imposes a restriction in the coding of the data because most cases of ser+resultant state are regarded as a passive voice. However, the evidence shows that ser is also functioning in cases where the resultant state is used just as an adjective and not a past participle (n=20/63; %≈31) (e.g., yo soy egresado de la carrera ‘I am an alumni of that major’ [h46]). This is evidence that the spread is not only of estar to the field once occupied exclusively by ser, but the spread also happens in the opposite direction (i.e., ser is spreading to places where once were exclusive to estar).

Discussion

From the previous review of the literature and description of the present study, three research questions were set forward. The first one was what linguistic and social variables help predict the use of estar and ser. The second was whether the variables that predict the use of estar were the same or different from the ones that predict the use of ser. And the third question was whether Spanish from highly educated Costa Ricans was different from those of speakers of the same age from other varieties of Spanish specifically in copula choice.

With regard to the first question, the variables that help predict the use of ser and estar, in Costa Rican Spanish, in the copula+adjective construction are adjective class, gradience, predicate reading, experience with the referent, and resultant state. Furthermore, the results of this study
show that, as demonstrated in previous studies, linguistic and semantic variables are the strongest predictors of copula choice in the structure copula + adjective. Concerning linguistic variables, the results show the following patterns: adjectives of size, physical appearance, mental/physical state, and status favor the use of estar (e.g., no es porque estén esteticamente lindas ‘it is not because they are aesthetically pretty’ [m30]) whereas adjectives of description/evaluation and description of a person(ality) disfavor it (e.g., el sentir de la mujer es importante ‘women’s feelings are important’ [m34]). Because of the unexpected results regarding gradience and the fact that they contradict previous theoretical accounts (i.e., Givón, 1984), a cross-tabulation of the variable gradience with each of the other independent variables was conducted to account for the behavior of this variable in the data.

When looking at gradience for ser and estar with regard to predicate reading, we find that estar and ser behave in a different manner. Figure 3 shows how gradient and absolute adjectives appear mostly with ser when the predicate has an individual-level reading (91% and 83% respectively) (e.g., uno [el sacerdote] que es un católico ‘one a priest that is a catholic’ [h30] and está muy tranquila ‘she is very calm’ [h30]). This finding contradicts Givón’s (1984) predictions that absolute adjectives will be found more in ser construction than in estar constructions and that gradient adjectives will pair more with estar because both absolute and gradient adjectives are found in similar percentages with ser and estar. In this case, we find partial support for the first part of this prediction because instead of just absolute adjectives we find that both absolute and gradient adjectives are found with ser almost categorically in predicates with an individual-level reading. Regarding the second part of the prediction, both absolute and gradient adjectives appear with both ser and estar. However when we analyze the adjectives separately in estar constructions, we find that the situation is not as clear-cut as with

ser constructions. In this case, gradient adjectives appear more with estar when the predicate has a stage-level reading (55.26%) (e.g., nada más estoy pendiente ‘I am only waiting for’ [h46]) whereas they appear to a lesser degree (44.74%) with estar when the predicate has an individual-level reading (e.g., el es exliceista ‘he is a Liceo alumnus’ [h30]). Absolute adjectives pattern in an opposite way to gradient adjectives with estar. Here we find that absolute adjectives appear more with estar if the predicate has an individual-level reading (i.e., a comparison of the individual to him/herself at a different point of time) (58.78%) (e.g., no estoy obligado a nadie ‘I am not obliged to anyone’ [h30] [in this case obliged means married] whereas if the predicate has a stage-level reading (i.e., a comparison of the individual to a group) they appear less with estar (47.22%) (e.g., católico es universal ‘catholic is universal’ [h28]). Again, these findings give partial support to Givón’s (1984) predictions and serve as empirical evidence to state that gradience is affected by the discursive variable: predicate reading. Even though these two variables are independent from one another, a deeper look at the properties of the discourse is needed to tease apart these differences with regard to the semantics of the adjective.

Figure 3 : Relationship between gradience and predicate reading
It also represents strong evidence that in native speaker data discursive variables are more important than semantic variables in the copula+adjective construction.

*Ongoing and indirect* experience with the referent disfavor the use of *estar* (e.g., *en el sentido de saber lo que es bueno* ‘in the sense of knowing what is good’ [m34]) whereas *immediate* experience with the referent favors it (e.g., *otros estaban pálidos* ‘others were pale’ [h37]). Finally, *non-resultant states* disfavor the use of *estar* (e.g., *como profesor fui muy jovial* ‘as a teacher I was very joyful’ [h46]) whereas *resultant-states* favor it (e.g., *que uno esta medio dormido* ‘that one is half asleep’ [h37]).

To answer question two, the evidence reveals that the same variables that help predict the use of *ser* help predict the use of *estar* and that *ser* and *estar* are in complementary distribution. The statistical analysis reveals that we have accurately defined the contexts where *ser* and *estar* are in variation since the same factors were found to be significant for both *ser* and *estar* as the application value. The analysis also shows that all variables were independent from one another and that there is no interaction among them. This evidence supports Geeslin’s and Díaz-Campos’ (2005) claim that we are dealing with a stable change and not a change in progress.

Regarding our third research question some of the results in this study are similar to those of previous studies regarding *copula+adjective* in Spanish. First, the variable *adjective class* pattern similarly to the results in Díaz-Campos and Geeslin (2005) group age 30-45, in that Spanish from Costa Rica shows a distribution of copula use where *estar* is favored with adjectives of *mental state* and *status*. However, in Costa Rican Spanish for the equivalent age group, *size and physical appearance* seem to pattern differently with *estar* being the preferred copula for these adjective types. With regard to adjectives of *description/evaluation* and *description of person(ality)* Caracas Spanish and Costa Rican Spanish pattern similarly with *ser* being the preferred copula and *estar* showing a high frequency that approaches a similar pattern to that of *ser* for the description of *person(ality)*. In all adjective types, we can see the presence of *estar* patterning similarly to *ser* and/or being the preferred copula.

After ruling out the interaction between *gradiency and predicate reading*, the second variable group is *predicate reading*. For this variable I coded for the *individual- and state-level* reading that is given to each predicate following Clements (2006). The variables are also similar to those found by Díaz-Campos and Geeslin (2005). *Stage-level readings* favor the use of *estar* while *individual-level readings* favor the use of *ser*.

The third variable group was *experience with the referent*. A pattern similar to that of Díaz-Campos and Geeslin (2005) is found for Costa Rican Spanish. *Immediate reactions* are almost categorically used with *estar* whereas *ongoing experience and indirect knowledge* seem to pattern evenly between the two copulas with *ser* still being the favored copula. In Díaz-Campos and Geeslin (2005), for the age group 30-45 there were no instances of *immediate reactions* whereas for ongoing experience *estar* was used almost categorically, and the instances of *indirect knowledge* were almost non-existent with *estar*. This discrepancy serves as empirical evidence that copula choice in Costa Rican Spanish is different to the choice in Venezuelan Spanish. However, because of the limited number of subjects of this study, evidence from a bigger sample are needed for claiming any structural different between these two dialects.

The fourth significant predictor group was *resultant state*. For this variable Díaz-Campos and Geeslin (2005) found that for the age group 30-45 this resultant state favors the use of *estar*. In the present study, a similar pattern emerged with resultant states favoring *estar*. However, the probability weights for *ser* and *estar* with respect to non-resultant states show that *estar* is being favor almost equally to *ser*. All other variable groups from Díaz-Campos and Geeslin (2005)
were not significant in the present study. These variables include frame of reference and susceptibility to change.

Conclusion

The variables such as adjective class, predicate reading, experience with the referent, and resultant state, have been established as strong predictors of copula choice in Spanish. In the present study, it has been demonstrated that these variables are also strong predictors of copula choice in Costa Rican Spanish. However, the model is not complete. The present study has demonstrated that by including other variables that are related to the semantics of the adjective, some predictors of previous models become non-significant. These variables that turned out to be not significant in the present model are Díaz-Campos and Geeslin’s (2004, 2005) Geeslin’s and Díaz-Campos’ (2005) susceptibility to change and frame of reference. Furthermore, the variables that were included in this study show evidence that the semantics of the adjective plays (i.e., adjective class and resultant state) an equally prominent role as discourse/pragmatics features (i.e., frame of reference, predicate reading, and experience with the referent) in first language Spanish. However, discourse/pragmatics variables are stronger predictors of copula choice. In the present study, a semantic feature (i.e., gradience) turned out to be a significant predictor of copula choice in Costa Rican Spanish. A deeper look at this variable shed light on the fact that the choice between ser and estar in Spanish can be found in discursive features as well as semantic features. Ser was used almost categorically to represent individual-level readings of the predicate. Estar shows more variation in this area. Gradient adjectives alternate with absolute adjectives with regard to the reading given to the predicate. For gradient adjectives if the predicate has an individual-level reading, ser is used more. If the predicate is given a stage-level reading, estar appears more times. However, for absolute adjectives if the predicate is given an individual-level reading, estar appears more times. If the predicate is given a stage-level reading, ser appears more times.

Further studies are needed to test the predictability of the new variables introduced in the present study as well as to include participants of the different age groups and social classes to study Costa Rican Spanish. Because the empirical evidence in this area has shown the power of discursive variables as well as semantic variables, more statistical model comparisons are needed in order to find the model that best fit the data regarding copula choice in copula+adjective constructions. The inclusion of subjects from different age groups will help overcome the major limitation of the present study which is the fact that only one age group was included, so any claim about change cannot be stated. More age groups and social classes will allow us to search for differences between varieties of Spanish with regard to copula choice and to find support of the hypothesis of change in progress (Silva-Corvalán, 1984) or of the hypothesis of stable change (Geeslin and Díaz-Campos, 2005). So far, this is not possible because this study can only be compared to parts of the previous studies. However, Díaz-Campos’ and Geeslin (2005) analysis of this phenomenon by age group and the similarity of the results with those found in the present study is evidence that we are moving in the right direction. Finally, evidence from this study shows that the spread is not only of estar to the field once occupied exclusively by ser, but that the spread also happens in the opposite direction (i.e., ser is spreading to places that were once exclusive to estar). This issue should be further studied in order to be able to explain and provide the field with a model of copula choice in Spanish.
References


