

# Syntactic and Pragmatic Usage of the [*estar* + Adjective] Construction in Puerto Rican Spanish: ¡Está brutal!

Esther L. Brown and Mayra Cortés-Torres  
University of Colorado-Boulder and Pima Community College

## 1. Introduction

The Spanish copulas *ser* and *estar* in combination with predicate adjectives have been the focus of many analyses in the study of Spanish grammar (Bolinger 1944, 1947; Bull 1942, 1965; Falk 1979a, 1979b, 1987; Luján 1981; Navas Ruiz 1963; Vañó-Cerdá 1982). Indeed, the synchronic variation and diachronic change of the copulas *ser* and *estar* in the context of predicate adjectives has been “one of the most thoroughly studied and most debated issues in Spanish linguistics” (Silva-Corvalán 1994:94). Considerable diachronic, geographic, social, and stylistic variation in the selection of one copula over the other (Guijarro-Fuentes & Geeslin 2006: 64) perhaps fuels the attention *ser/estar*-usage receives in the literature. Previous analyses identify multiple linguistic and extralinguistic factors that play a role in determining copula choice, and it is generally no longer sustained that an identifiable list of semantic attributes (e.g.; permanent vs. temporary, inherent vs. attributive) can unilaterally predict the selection of one copula over another. As Díaz-Campos & Geeslin (2011:75) summarize, previous research has shown us that “the variation in use of *ser* and *estar* has been ongoing for a long time, that not all adjectival contexts are affected in the same way and at the same time, and that the extension of the use of *estar* has occurred across time”.

This present work provides a new perspective from which to examine [copula + adjective] occurrences by employing a usage-based approach in our analyses. We examine variable copula selection in conversational Puerto Rican Spanish; a choice which is often not predictable from the adjective, as can be seen in examples (1) and (2):

(1) ...sí **son brutales**... (Interview 8, page 2, speaker 10)

(2) Y esas son las playas que **están brutales** (Interview 2, page 22, speaker 2)

We propose that [*ser* + adjective] and [*estar* + adjective] are constructions (discussed below) and highlight how particular instances of these constructions have acquired specific pragmatic and semantic characteristics (Bybee 2006). Further, the usage-based approach we adopt (Bybee 2001) enables us to propose specific mechanisms and pathways of change that account for the expanding range of uses of *estar* at the expense of *ser*.

## 2. Background

The contexts of use of [*estar* + adjective] have expanded in the development from Latin to contemporary Spanish (Vañó-Cerdá 1982). *Estar* is now used with adjectives previously only co-occurring with *ser*, and this semantic extension is apparent in both monolingual and bilingual communities (Behrend 1986; De Jonge 1993a, 1993b; Gutiérrez 1992, 1994; Malaver 2000; Ortiz

---

\* We would like to acknowledge the helpful comments we received from three anonymous reviewers and constructive feedback from the audience attending the Workshop on Caribbean Spanish at the Hispanic Linguistics Symposium 2010.

López 2000; Silva Corvalán 1986, 1994). Recent work on copula choice (e.g.; Cortés-Torres 2004; Díaz-Campos and Geeslin 2011; Geeslin and Guijarro-Fuentes 2007a, 2007b, 2008; Guijarro-Fuentes and Geeslin 2006) employs variationist methodologies and brings to light many linguistic (resultant state, adjective class, predicate type, experience with the referent, frame of reference, susceptibility to change), extralinguistic (socioeconomic level, age, level of educational attainment), and stylistic factors that significantly constrain the *ser/estar* variation across different varieties of Spanish. Together these works identify strong probabilistic tendencies for selection of *ser* or *estar* with adjectives, and illustrate that in the majority of the instances of use, copula selection in Spanish is predictable through identification of context specific (extra)linguistic variables.

Ortiz López (2000) has, to our knowledge, conducted the only study on copula choice variation in Puerto Rican Spanish. In his study he found that *estar* was used in contexts normally associated with *ser* in 19% of the answers to a questionnaire and in 12% in oral interviews.<sup>1</sup> Traditionally, these cases of semantic extension into contexts prescriptively reserved for *ser* are deemed *innovative estar* uses (Gutiérrez 1992, 1994; Ortiz López 2000; Silva-Corvalán 1994, 2001). These innovative uses will be a primary focus of this paper. We argue that innovative uses are precisely instances of [*estar* + adjective] that fall outside of the syntactically predictable uses of *estar* as identified in this, and in previous, work. Additionally, we explore whether there is pragmatic meaning attached to these innovative uses and if we can give an explanation for the directionality of this diachronic change (that is the diachronic preference for more contexts for *estar* as opposed to *ser*).

Building upon work conducted on other varieties of Spanish (Cortés-Torres 2004; Díaz-Campos and Geeslin 2011; Geeslin and Guijarro-Fuentes 2007a, 2007b, 2008; Guijarro-Fuentes and Geeslin 2006), we employ variationist methodology to identify the linguistic factors that favor *ser* and *estar* in the [copula + adjective] context in Puerto Rican spontaneous discourse. Based on the patterns of use we identify, we propose an analysis that considers [*ser* + adjective] and [*estar* + adjective] as constructions. According to Bybee & Eddington (2006), a string of words or morphemes that is used with some frequency can be considered a construction even if its form, function and/or meaning are entirely predictable. Such strings of words used frequently together constitute items in cognitive storage, representations to which changes in function or form can be attached (Bybee 2001, 2006). As Bybee & Torres Cacoullos (2008: 409) note, “constructions – form meaning pairings – are usage-based; that is, they represent specific lexical items and their social and pragmatic contexts of use”, and further, constructions can “often take on non-transparent meanings” (Bybee 2001:343).

We argue that the non-prototypical (‘innovative’) uses of [*estar* + adjective] express a pragmatic meaning of the construction as a whole. The pragmatic meaning of the construction is independent from the meaning of the adjective and arises from speakers’ understanding and interpretation of how the [copula + adjective] constructions are used generally in discourse (as determined by our variable rule analysis). That is, the pragmatic meaning is derived and acquired through use. In addition, in a discussion that sets our study apart from previous analyses, we look at the potential source or explanation for the productive *estar* uses by focusing on the specific adjectives that appear to be driving this use and identify their lexical characteristics using semantic clusters. The following section outlines the data and methods we employ to substantiate these claims.

### 3. Data & Methods

Usage-based approaches to language hold that real data are imperative to understanding patterns of variation as well as the timing and origins of linguistic change (Boyland 2010). Thus, in our investigation of the syntactic and pragmatic uses of [copula + adjective] constructions in Puerto Rican Spanish, we utilized a corpus of recorded, spontaneous conversations between native speakers of Puerto Rico. Conversations ranged in duration from one half hour to three hours in length, and were all conducted, recorded, and transcribed in 2002 as part of a different study (Cortés-Torres 2005). The informal conversations total approximately 370,000 words of spoken Puerto Rican Spanish representing roughly 27 hours of conversation between the interviewer and 31 native-speaking close

---

<sup>1</sup> The adjectives that showed a higher use of *estar* in this context were related to state, age, size, and class. Further, men used *estar* more than women and monolinguals more than bilinguals.

friends and family members. We extracted from the corpus all instances of *ser* and *estar* used with a predicate adjective<sup>2</sup>. We included cases with and without intervening material, as illustrated in (3) and (4) respectively.

(3) porque **es bien chavona** para eso (Interview 11, page 23, speaker 18)

(4) La casa **está chévere** (Interview 9, page 7, speaker 15)

This gave us a total of 2566 tokens of *ser* and *estar* + adjective for analysis. In order to isolate as objectively as possible instances of innovative (or non-syntactically predicted) uses of *estar*, we coded all tokens for linguistic and extralinguistic variables identified as significantly predicting *ser* or *estar* in other varieties of Spanish in order to determine if the same holds true for Puerto Rican Spanish. Following Geeslin & Guijarro-Fuentes (2008) and Díaz-Campos and Geeslin (2011), therefore, we coded for the following linguistic variables. (The Appendix lists the number of tokens for each factor.)

1. The dependent variable (*ser*, *estar*): whether a form of the verb *ser* or *estar* was used.
2. Susceptibility to change (yes, no): This factor group contrasts characteristics that are changeable or susceptible to change (such as the length of one's hair) with those that are not (such as the length of a room).
3. The frame of reference (class, individual): Class frame refers to cases in which the referent is compared to a group of like entities, while individual frame describes instances in which a referent is compared to itself at a different point in time. The inherent difficulty of coding this factor group in naturalistic data was lessened by the interviewer's (Cortés-Torres') shared knowledge of the people and places commented upon in the conversations.
4. Experience with the referent (indirect/ongoing, immediate): This factor group differentiates "statements that are dependent on the experience of the speaker from those that do not stem from direct experience with the referent" (Díaz-Campos & Geeslin 2011:77). Coding experience with the referent is encumbered by the need to know (or intuit) each speaker's relationship with the referent. Thus we created a binary division between indirect/ongoing experiences contrasted with "cases where a single event or moment in time is described" (Díaz-Campos & Geeslin 2011:77).
5. Adjective class (age, size, evaluation/description, physical or mental state): This factor group describes the adjectives themselves and closely reflects categories identified by Geeslin & Guijarro-Fuentes (2008). We were not able to code confidently a distinction between evaluative and descriptive comments, and therefore considered evaluation/description as one factor within this group.
6. Resultant state (yes, no): This independent variable codes for the potential effect of the perfective aspect reflected in adjectives that denote a resultant state of a verb (*alterado*) compared to those adjectives which do not (*amarillo*).

Following Ortiz (2000), we code for the following extralinguistic variables based upon speaker information. The distribution of speakers across extralinguistic factor groups is summarized in Table 1.

1. Gender (male, female): Speakers were coded according to their biological gender. We analyze the speech of fourteen men and seventeen women.
2. Degree of bilingualism (monolingual, bilingual); those speaking only Spanish were classified as monolingual, and those who used English at work, but Spanish at home, as well as those who used both languages interchangeably were coded as bilingual. The speech of twenty monolinguals and eleven bilinguals was used in this analysis.
3. Age (20-29, 30-59, 60+): We divided the speakers in age groups representing different stages of life. This method, known as *emic*, classifies speakers according to shared time experiences that might correlate with life states or their history (Eckert 1997:155).

---

<sup>2</sup> We did not extract the invariable cases of [*ser* + adjective] reflecting nationality or religion (N = 67). These adjective classes represent the only exclusion.

**Table 1: Summary of speakers numbers for extralinguistic variables.**

	Monolingual	Bilingual
<b>Age 20-29</b>		
Male	3	2
Female	3	0
<b>Age 30-59</b>		
Male	2	2
Female	3	6
<b>Age 60+</b>		
Male	4	1
Female	4	1

## 4. Results

We examine all 2566 cases of copula + adjective. In this Puerto Rican data, 39% of the cases are a form of [*estar* + adjective], with the bulk of the data representing *ser* tokens. We find 887 different adjective types including paradigmatically related forms (masculine, feminine, singular, plural). If we conflate these forms into lemmas, however (*bueno*, *bueno*, *buenos*, *buenas* > *bueno*), we find 600 different adjectives used in our corpus with the copulas. Table 2 summarizes the usage patterns of the copulas with predicate adjectives as occurring in our data.

**Table 2: Distribution of *estar* and *ser* with adjective lemmas**

	Lemmas	Tokens	% Lemmas	Example
Only <i>ser</i>	315	626	53	<i>es fácil</i>
Only <i>estar</i>	187	435	31	<i>está acostumbrado</i>
Variable: <i>ser</i> or <i>estar</i>	98	1505	<b>16</b>	<i>es/está bueno</i>
Total:	600	2566	100	

Of these 600 adjectives, 315 pertain exclusively to *ser*. There are 626 tokens of invariable [*ser* + adjective] use. The most frequent example of an adjective used exclusively with *ser* is *fácil* (N = 26). We have 187 adjectives that are used exclusively with *estar* (and 435 tokens of these adjectives). The most frequent adjective used exclusively with *estar* is *acostumbrado* (N = 22). Of the 600 adjective lemmas, 98 are used with both *ser* and *estar*, the most frequent of which is *bueno* (N = 222). The most frequent variable adjectives in the Puerto Rican discourse (those with more than 20 tokens each) are summarized in Table 3. The percent of use with *estar* varies from the near exclusive use with *estar* for *brutal* (96%), to the near exclusive use of *ser* with *diferente* (2%).

**Table 3: Variable adjectives (lemmas) >20 tokens**

Adjective	% <i>estar</i>	N	Adjective	% <i>estar</i>	N
brutal	<b>96</b>	24	lindo	<b>32</b>	28
seguro	<b>79</b>	28	igual	<b>31</b>	39
flaco	<b>68</b>	22	caro	<b>28</b>	29
gordo	<b>68</b>	22	raro	<b>26</b>	23
rico	<b>62</b>	26	fuerte	<b>24</b>	34
loco	<b>61</b>	54	grande	<b>21</b>	92
malo	<b>53</b>	64	chiquito	<b>12</b>	33
tranquilo	<b>45</b>	29	difícil	<b>8</b>	40
bonito	<b>45</b>	96	nuevo	<b>5</b>	21
bueno	<b>34</b>	222	diferente	<b>2</b>	66

These 98 variable adjectives make up 16% of the adjectives in our corpus, with 84% of the adjectives, therefore, invariably selecting one copula over the other. It is interesting to note that by some accounts (Vañó-Cerdá 1982) as many as 80 percent or more of the adjectives in Spanish allow *ser/estar* contrast. Clearly, as seen in Table 2, the 16% in our data does not approximate this 80% figure. This discrepancy between what is grammatically permissible and what is appearing in actual language use highlights, in our view, an important advantage of corpus-based analyses. That is, corpus data may more faithfully depict the characteristics of the input that play a role in language acquisition and adult language variation.

In order to ascertain which linguistic and extralinguistic factors constrain copula use in spontaneous Puerto Rican discourse, we conducted a regression analysis using Rbrul (Johnson 2009) on the tokens that we extracted from the recorded conversations. The application value is use of *estar* and all the linguistic and extralinguistic factors we summarized previously are included in the analysis. Results of our regression analysis are summarized in Table 4.

**Table 4: Linguistic and extralinguistic factors conditioning copula use in Puerto Rican discourse**

factor	logodds	tokens	proportion	centered weight	
<b>Frame of reference</b>					
individual	2.12	680	0.976	0.893	
class	-2.12	1886	0.182	0.107	
<b>Experience with the referent</b>					
immediate	1.722	632	0.853	0.848	
ongoing	-0.750	1520	0.223	0.321	
indirect	-0.972	414	0.312	0.274	
<b>Resultant state</b>					
resultant	1.022	417	0.849	0.735	
non-resultant	-1.022	2149	0.304	0.265	
<b>Adjective class</b>					
mental state	1.633	206	0.840	0.837	
physical state	1.096	219	0.922	0.75	
evaluation/description	-0.438	1936	0.310	0.392	
age	-0.787	86	0.233	0.313	
size	-1.504	119	0.092	0.182	
<b>Age</b>					
20-29	0.330	613	0.361	0.582	
30-59	0.019	1353	0.402	0.505	
60+	-0.349	600	0.403	0.414	
<b>deviance</b>	<b>df</b>	<b>intercept</b>	<b>total</b>	<b>mean</b>	<b>input prob.</b>
1487.689	11	2.027	2566	0.392	0.884

The factors selected as significant are listed in Table 4 in order of their magnitude of effect. The linguistic factor group with the greatest magnitude of effect is the frame of reference. When the frame of reference is the individual (that is, comparing an individual to him or herself), the use of *estar* is likely (logodds 2.12). Conversely, when the frame of reference is a class (as opposed to the individual), use of *estar* is disfavored (logodds -2.12). This result is in line with previous analyses (Silva-Corvalán 1986, 1994; Gutiérrez 1992, 1994; Cortés-Torres 2004; Díaz-Campos & Geeslin 2011).

Also selected as significant is the speaker's experience with the referent. If the experience is immediate, use of *estar* is more likely than *ser* (logodds 1.722). When the speaker has ongoing or indirect experience with the referent, however, *estar* is disfavored. This result concurs with what Guijarro-Fuentes & Geeslin (2006), Geeslin & Guijarro-Fuentes (2007a, 2007b, 2008) and Díaz-Campos & Geeslin (2011) found in their studies.

Resultant state adjectives also predict *estar* usage. Díaz-Campos & Geeslin (2011) determined that resultant state adjectives (acostumbrado, errado) all strongly favored *estar* in the Venezuelan variety

they studied. Conversely, in the Puerto Rican data, adjectives not reflecting a verbal resultant state do not predict *estar* usage (logodds -1.022).

Also in line with previous research (Silva-Corvalán 1986, 1994; Gutiérrez 1992, 1994; Ortíz López 2000; Cortés-Torres 2004; Guijarro-Fuentes & Geeslin 2006; Geeslin & Guijarro-Fuentes 2007a, 2007b, 2008; Díaz-Campos & Geeslin 2011) adjective class is a significant predictor of copula use. When the adjective describes a mental or physical state, use of *estar* is favored. When the adjective belongs to the class of evaluation, description, age or size, *estar* is disfavored.

The linguistic factor group susceptibility to change was not selected as significant. This factor group was selected as significant in other studies, although it was not significant in the older population studied by Díaz-Campos & Geeslin (2011).

The only extralinguistic factor group selected as making a significant contribution to the model of copula use in Puerto Rican spontaneous discourse was age. The only group to disfavor *estar* usage (logodds -0.349) was the oldest generation (60+). Both groups of younger speakers (20-29, 30-59) favor *estar* usage. We are not currently able to adjudicate between possible explanations for this age pattern (i.e.; whether this reflects a change in progress or age grading). Not selected as significant are gender and degree of bilingualism. Given previous studies on other varieties of Spanish [e.g.; Gutiérrez (1992, 1994), Silva-Corvalán (1986, 1994), Guijarro-Fuentes & Geeslin 2006, Geeslin & Guijarro-Fuentes 2007a, 2007b, 2008], the lack of significant result for degree of bilingualism is noteworthy and perhaps deserves further research.

With the results we summarize in Table 4, the *ser/estar* variation is better understood. To illustrate, given the syntactic and lexical factors we discussed, the copula use with the variable adjective *gordo* in examples (5) and (6) is highly probable. In example (5) the speaker is contrasting the new physical state of individual “C” with his previous form (individual frame of reference) and is surprised by the weight change (immediate experience), whereas in example (6), the speaker has ongoing experience with the referent and is comparing her to the more general population (or class).

(5) Pero **está más gordo**, C. está más gordo (Interview 15, page 7, speaker 25)

(6) Pues ella era gordita, **es gordita**. (Interview 4, page 42, speaker 4)

The linguistic factors that predict *ser* or *estar* use, as indicated in the results of the Rbrul analysis, reflect very strong tendencies. That is, forms of *ser* and *estar* each frequently co-occur in use with these specific lexico-syntactic functions. The usage-based framework we adopt, “emphasizes the role of frequency, or repetition, in the formation of what we call grammar” (Scheibman 2001:63), and assumes that speakers are sensitive to and have cognitive representations that reflect these patterns of use (Bybee 2006). Indeed, within exemplar models, grammar is considered to be the cognitive organization that reflects speakers’ experiences with language (Bybee 2006:711). Thus, our findings suggest the following prototypical grammars for [*ser* + adjective] and [*estar* + adjective] in Puerto Rican conversational data, summarized in Table 5.

**Table 5: Prototypical use for *ser/estar* + adjective in PR**

Canonical use:	Frame of reference	Adjective class	Resultant State	Experience with referent
[ <i>ser</i> + adjective]	class	age, size	non-resultant	ongoing, indirect
[ <i>estar</i> + adjective]	individual	mental, physical state	resultant	immediate

This ‘grammar’ reflected in these linguistic factors, in fact, describes and accounts for the vast majority of the 2566 uses of [*ser/estar* + adjective] that we examine. It does not account for all tokens, however. We have 116 [*estar* + adjective] tokens that are not explainable by these linguistic factors. That is, if we exclude all [*estar* + adjective] cases that are described in Table 5, those where *estar* is used with the individual and immediate reference with an adjective describing a mental, physical, or

resultant state, we are left with ‘innovative’ uses of *estar*. We will be classifying innovative uses as pragmatic uses based upon the meaning we propose the construction may acquire.

The instances of [*estar* + adjective] not predictable from linguistic or lexical factors (as determined by our regression analysis) total 116 cases, or 5% of our data. An example can be seen in (7). To clarify, (7) illustrates the cases of [*estar* + adjective] for frame of reference ‘class’, experience with referent ‘ongoing or indirect’, and adjective classes other than ‘mental’, ‘physical’ or ‘resultant state’.

- (7) E: O sea, ¿los martes todo cierra temprano?  
 I(Q): Por lo menos esa parte ahí sí.  
 I(N): No, sí pero siempre hay-  
 I(Q): No, pero, pero el Pulpo loco, Pulpo lo-, el Pulpo loco donde fuimos el año pasado ahí no, eso está abierto. Pero lo, Soleil ya como a las ocho cierran.  
 I(N): Pero siempre hay *spots* por ahí que uno puede ir.  
 I(Q): Sí, podemos ir a otro sitio, bueno ahí **está chévere**, ahí **está relax**, la playa.  
 I(N): Ay, el Viejo San Juan que hace tiempo yo no voy al Viejo San Juan pa' estar allí en una barra o algo.  
 (Interview 2, pages 8-9, speakers 2 & 3)

For the remainder of this work we limit our focus and discussion to these 116 pragmatic [*estar* + adjective] cases: instances in which the linguistic factors do not predict *estar* usage, but rather, based on the syntactic and lexical factors, one would predict *ser*.

What is the meaning of this construction? We suggest that the construction [*estar* + adjective] can acquire meaning from the prototypical uses that we described in Table 5 (immediacy, individual frame). That is, speakers, through repetition, come to associate these characteristics with the [*estar* + adjective] construction. Since *estar* is the strongly preferred copula in a probabilistic sense for cases of immediacy and individual, this meaning becomes attached to the construction. The generalized implication of immediacy, often accompanied by individual, in other words, is correlated with this construction and stems from multiple instances of use with this meaning. Bybee & Torres Cacoullós (2008:410) summarize how such form meaning correlations are tracked:

Usage-based grammar, coupled with the rich memory representations of exemplar models, allows a direct representation of variation and change in cognitive representations. Because it tracks frequency of co-occurrence, this model provides us with ways of explaining how new constructions arise gradually over time, how specific instances of constructions contribute to more general change and how specific instances take on special value, be it semantic or social.

We speculate that this can provide a partial explanation for the preference of *estar* over *ser*, particularly in spontaneous, conversational settings, such as the data we analyze. The selection of non-syntactically prescribed *estar* over *ser* carries with it the immediate/individual sense of the construction, and hence adds to the pragmatic weight of the utterance (perhaps through increased foregrounding). This interpretation might find partial parallel in Hopper & Thompson (1980). On that study of transitivity, it is found that both immediacy and individuation add to the transitivity of a construction. Increased transitivity means increased foregroundedness. We do not wish to argue that *estar* is transitive, but rather that the construction carries with it implications associated with foregrounded information, and hence perhaps more pragmatic weight. This is a similar view to Fernández Leborans (1999:2366-67) who argues that by using *estar* with adjectives the speaker is portraying a subjective view of the situation. Silva Corvalán (1994: 108) has also argued that the ‘absolute frequency of occurrence of *estar* with certain individual attributes may be promoting its extension to less frequent class frame contexts in which the same attributes are used’, explaining the use of *estar* with certain adjectives in non-syntactically prescribed uses. Similarly, Scheibman (2001: 65) suggests a general hypothesis regarding linguistic constructions that predicts an expectation of a

“greater co-occurrence of elements whose combinations lend themselves to conveying speaker point of view than those whose combinations do not...”. In this way *estar*, which reflects a speaker’s more immediate experience with a referent, has an advantage over *ser*.

In their study of adjective use following verbs of ‘becoming’, Bybee and Eddington (2006) found evidence that high-frequency lexical instances of [verb + adjective] constructions act as the central members of categories and are essentially prefabs [prefabricated or conventionalized word sequences (Bybee 2006:721)]. To examine these categories, Bybee and Eddington (2006) group adjectives, associated through use with specific verbs in the construction, into semantically related clusters or clouds. As Vergara Wilson (2009:277) notes, specific [verb + adjective] combinations “with high token frequency serve as central members of these categories and [can] be thought of as prototypical, especially since they tend to display most of the features common in other members”. Further, within these clusters of semantically related adjectives, “there is ‘family resemblance’ where marginal members may share characteristics with the central members but not necessarily with one another” (Vergara Wilson 2009:277). Less frequent adjectives that share features with the frequent central members become associated with the construction through their semantic relatedness to the central member and are used productively in the open adjective slot of the construction. Their productive uses, therefore, reflect cognitive associations with central members of the category.

In the 116 pragmatic tokens in our data there are 47 different adjective lemmas used pragmatically. What is noteworthy, however, is that over half the tokens (N = 60) come from just 7 adjectives, summarized in Table 6.<sup>3</sup> This suggests that the use of pragmatic [*estar* + adjective] is heavily dominated by formulaic uses or fixed phrases such as *está brutal*.

**Table 6: Most frequent adjective lemmas with pragmatic *estar* use**

Adjective	N
brutal	16
bueno	12
bonito	12
malo	5
fuerte	5
chévere	5
caro	5
Total:	60

It seems that the most frequent adjectives *brutal*, *bueno*, and *bonito* are, in fact, acting as anchors or central members in this pragmatic use. It might be argued that *brutal* has become conventionalized as a fixed expression given its high frequency use with *estar* (96%) (Bybee & Eddington 2006). What accounts for the other 48% of the adjectives used with this pragmatic function? What motivates their use?

We examined all the adjectives used with a pragmatic function in our corpus and placed them in semantically related clusters. This was done based on their intended meaning as used in the corpus (e.g.; *sucio* as ‘bad’ as opposed to ‘dirty’) and based on one of the author’s native speaker intuition (Cortés-Torres). In fact, if we just consider the semantic clusters of the three most frequent adjectives in Table 6, the vast majority of the pragmatic uses (86%) are accounted for as can be seen in the following illustrations.

In the first illustration, we have a particularly strong cluster with three closely related adjectives *brutal*<sup>4</sup>, *malo*, and *fuerte* all used with relatively high frequency with a pragmatic function. These three adjectives alone account for nearly 20% of the pragmatic uses. This semantic cluster, with high frequency and close semantic relationships, forms a strong central category. This explains the

<sup>3</sup> The total N for these adjectives reflects their use solely in the pragmatic [*estar* + adjective] construction and not their use in the corpus overall.

<sup>4</sup> The adjective *brutal* is used with both a positive (N = 11) and a negative (N = 5) connotation in this corpus. When used to denote positive attributes, this adjective would be contributing to the cluster illustrated in Figure 2, *estar bueno*.

productive nature of the category evidenced in this illustration by the many types of adjectives used. Many singleton adjectives (*difícil*, *cabrón*, *bestial*, *terrible*, *etc.*), related semantically to the central members, are used with this pragmatic function, and this cluster accounts for 39% of *estar* with pragmatic uses. This information is summarized in Table 7 following the figures.

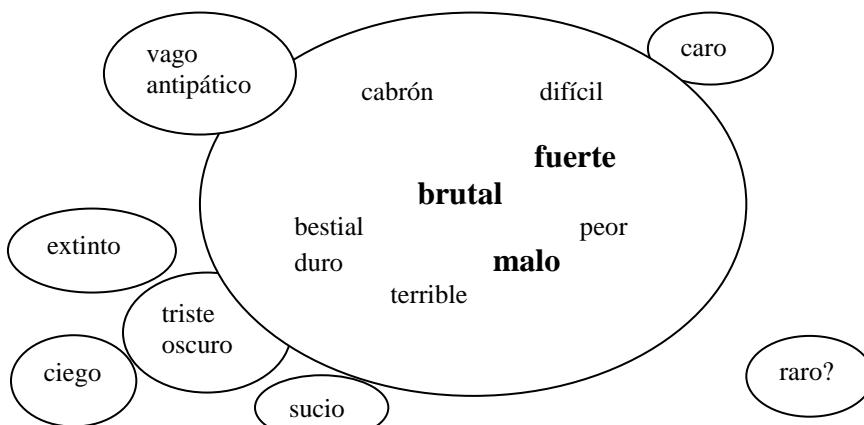


Figure 1. Cluster centering on *estar brutal*

Additionally, we propose a cluster based on *bueno* and simultaneously *chévere* and *brutal* (with positive connotations), which is also relatively productive. This cluster accounts for 31% of the pragmatic cases as can also be seen in Table 7.

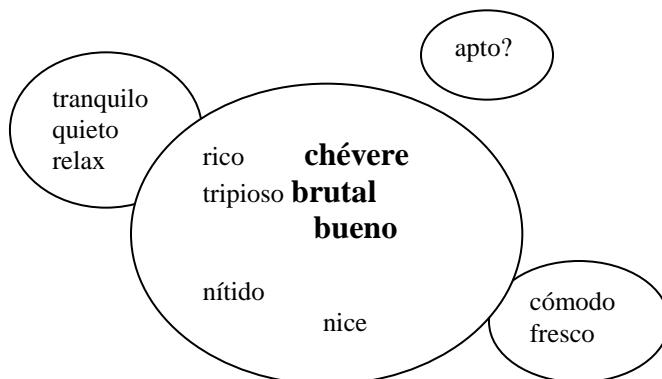


Figure 2. Clustering centering on *estar bueno*

In the third illustration (Figure 3) we see the adjective 'bonito' at the center of the cluster, an adjective frequently used with the pragmatic meaning (12 times in our data). *Estar bonito* as a construction can form a lexical representation. Then other semantically related adjectives can be used productively in this construction through analogy to this central member. Yet again, the suggestion is that the pragmatic uses of the other adjectives (such as *lindo* and *precioso* in this case) are motivated by their semantic relatedness to the central member of the cluster.

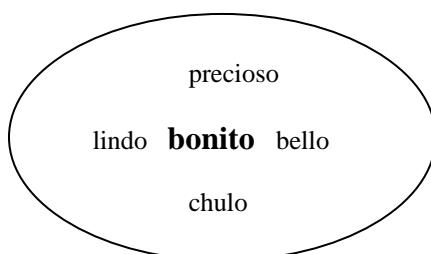


Figure 3. Cluster centering on *estar bonito*

Lastly, Figure 4 presents a group of adjectives with very close semantic relationships pertaining to size. The central member may be *chiquito* with the highest token frequency of the group (3), as is summarized in Table 7.



Figure 4. Cluster centering on *estar chiquito*

In Table 7 it is apparent that 95% of the innovative uses stem from 4 clusters of semantically related adjectives (Figures 1-4). Thus, as argued by Bybee & Eddington (2006), we claim that the speaker's choice of *estar* in the 116 pragmatic uses reflects his/her accumulated experience with the use of the construction generally and the productive uses reflect analogical extensions to central category members.

**Table 7: Clusters of pragmatic *estar* usage in Puerto Rican discourse**

	Adjectives	N	% Total Pragmatic Uses
<b>Figure 1: brutal</b>	brutal	11	
	fuerte	5	
	malo	5	
	caro	5	
	cabrón	2	
	peor	2	
	raro	2	
	ciego	2	
	triste	2	
	bestial, terrible, difícil, duro, sucio,	9	
	vago, antipático, extinto, oscuro		
	Total	<b>45</b>	<b>39</b>
<b>Figure 2: bueno</b>	bueno	12	
	chévere	5	
	brutal	5	
	relax	2	
	quieto	2	
	rico	2	
	nice	2	
	cómodo	2	
	nítido, tripioso, apto, tranquilo	4	
		Total	<b>36</b>

<b>Figure 3: bonito</b>	bonito	12	
	chulo	2	
	lindo	2	
	precioso	2	
	bello	1	
	Total	<b>19</b>	<b>16</b>
<b>Figure 4: chiquito</b>	chiquito	3	
	pequeño, mediano, alto, grande, gordo, grueso, estrecho	7	
	Total	<b>10</b>	<b>9</b>
Other adjectives	Completo (2), claro, fresco, <i>full-time</i> , joven	<b>6</b>	<b>5</b>
Total		<b>116</b>	<b>100</b>

## 5. Discussion & Conclusion

In this work we examined [copula + adjective] use in naturally occurring spontaneous Puerto Rican discourse to contribute to the growing body of literature on innovative uses of *estar*. Of the 2566 cases of [copula + adjective] examined, 39% of the cases are a form of [*estar* + adjective], with the bulk of the data representing *ser* tokens. The factors that were selected as predictors of *estar* in the multiple regression analysis were frame of reference (individual), experience with the referent (immediate), whether the adjective reflected a verbal resultant state (resultant state), adjective class (physical/mental state), and age of the speaker (speakers under the age of 60). However, there were 116 cases of [*estar* + adjective] used in our data that could not be accounted for using the prototypical meanings derived from the linguistic factors identified using Rbrul. We labeled these ‘pragmatic’ cases and argued that [*estar* + adjective] is a construction with a separate meaning ascribed to it through repeated patterns of use in discourse. The selection of *estar* over *ser* carries with it the immediate/individual sense of the construction, and hence adds to the pragmatic weight of the utterance. We looked at all the adjectives used with a pragmatic sense and placed them in semantic clouds that might explain their use with this sense by their relatedness to the central members of the clusters.

The results we report are in line with the usage-based view adopted by Bybee (2006:730) that “grammar is built up from specific instances of use that marry lexical items with constructions; it is routinized and entrenched by repetition and schematized by the categorization of exemplars”. By viewing [*estar* + adjective] and [*ser* + adjective] variation as instantiations of speakers’ cognitive organizations reflective of use, we are able to provide a description of not just the variable *ser/estar* patterns in Puerto Rican discourse, but we are also potentially able to predict and explain directions of change based on frequency of use and similarity to central category members. Specifically, the pragmatic [*estar* + adjective] uses are argued to be derived from analogical extensions to previous experiences speakers have with the construction (Bybee & Eddington 2006). By looking at recurrent patterns of use in the spontaneous discourse, we may be able to test the potential role of constructions in variation and change and to know which constructions speakers are using and storing.

We were unable to test with our current corpus the diachronic trajectory of this variation in Puerto Rican Spanish. Research suggests, however, that analyses of both distributional patterns and changes in relative frequency measures can give us a more precise measure of the strength of the constructions we propose (Torres Cacoullós and Walker in press). Thus, in future studies we would like to consider adjacency (the existence of intervening material or not within the [*estar* + adjective] construction) and the strength of association (adjectives variably or invariably co-occurring with *estar*) for the pragmatic uses both synchronically, and with appropriate data, diachronically. Future analyses should also consider other extralinguistic variables not quantified in this analysis (socioeconomic status, level of education, etc.). Additionally, we would like to devise a way to conduct an analysis on all 2566 of our adjectives similar to the one we reserved for the 116 pragmatic uses in this study. Such an analysis, as

noted by an anonymous reviewer, might reveal possible meaning changes in the adjectives when used in constructions contrasted with the adjective's basic meaning(s), thus shedding more light on how the [*estar* + adjective] construction functions.

## Appendix

### Percentage of *estar* use for linguistic and extralinguistic variables

		<i>Ser</i>	<i>Estar</i>	Total N	% <i>Estar</i>
		N	N		
<b>Susceptibility to Change</b>	Yes	1184	927	2111	44
	No	375	80	455	18
<b>Frame of Reference</b>	Individual	16	664	680	98
	Class	1543	343	1886	18
<b>Experience</b>	Immediate	93	539	632	85
	Indirect/ongoing	1466	468	1934	24
<b>Adjective class</b>	Physical/mental state	50	375	425	88
	Evaluation/description	1335	601	1936	31
	Age	66	20	66	30
	Size	108	11	119	9
<b>Resultant State</b>	Yes	63	353	416	85
	No	1496	654	2149	30
<b>Gender</b>	Female	922	605	1527	40
	Male	637	402	1039	39
<b>Language(s)</b>	Bilingual	866	567	1433	40
	Monolingual	693	440	1133	39
<b>Age</b>	60+	358	242	600	40
	30-59	809	544	1353	40
	20-29	392	221	613	36
<b>Total</b>		1559	1007	2566	39

## References

- Behrend, Elvira H. 1986. The use of *Ser* and *Estar* by Bilingual Mexican-Americans in the Chicago area: A Language in contact study. Hamburg: University of Hamburg thesis.
- Bolinger, Dwight. 1944. More on “*Ser*” and “*Estar*”. *The Modern Language Journal* 28.233-38.
- 1947. Still more on *Ser* and *Estar*. *Hispania* 30.361-67.
- Boyland, Joyce T. 2010. Usage-based models of language. *Experimental and Quantitative Linguistics*, ed. by David Eddington, 351-419. Munich: Lincom.
- Bull, William E. 1942. New principles for some Spanish equivalents of “To Be”. *Hispania* 25.433-43.
- Bull, William E. 1965. *Spanish for Teachers*. New York: Ronald Press Co.
- Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press
- . 2006. From usage to grammar: The mind’s response to repetition. *Language* 82.711-33.
- and David Eddington. 2006. A usage-based approach to Spanish verbs of ‘becoming’. *Language* 82.323-55.
- and Rena Torres Cacoullous. 2008. Phonological and Grammatical Variation in Exemplar Models. *Studies in Hispanic Lusophone Linguistics* 1. 399-413.
- 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.

- Cortés-Torres, Mayra. 2005. La perífrasis de estar + -ndo en el español puertorriqueño: ¿variación dialectal o contacto lingüístico? Albuquerque, NM: University of New Mexico dissertation.
- De Jonge, Bob. 1993a. (Dis)continuity in language change: ser and estar + age in Latin American Spanish. *Linguistics in the Netherlands*, ed. by Frank Driekoningen and Kees Hengeveld, 69-80. Amsterdam: Benjamins.
- 1993b. Pragmatismo y gramaticalización en el cambio lingüístico: ser y edad en expresiones de edad. *Nueva Revista de Filología Hispánica* 41.99-126.
- 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.
- Eckert, Penelope. 1997. Age as a sociolinguistic variable. *The handbook of sociolinguistics*. ed. by Florian Coulmas, 151-67. Oxford: Blackwell.
- Falk, Johan. 1979a. Ser y estar con atributos adjetivales. Stockholm: Almqvist & Wiksell International.
- 1979b. Visión de norma general versus vision de norma individual. *Studia Neophilologica* 51.275-93.
- 1987. Reflexiones en torno a Vañó-Cerdá, ser y estar + adjetivos. *Studia Neophilologica* 59.109-28.
- Fernández Leborans, María Jesús. 1999. La predicación: Las oraciones copulativas. *Gramática descriptiva de la lengua española*, ed. by Ignacio Bosque Muñoz and Violeta Demonte Barreto, 2354-2460. Madrid: Espasa Calpe.
- Geeslin, Kimberly, and Pedro Guijarro-Fuentes. 2007a. Exploring Copula Choice in Spanish: A Look at Gender. *Selected Proceedings of the Third Workshop on Spanish Sociolinguistics*, ed. by Jonathan Holmquist et al., 92-102. Somerville, MA: Cascadilla Proceedings Project.
- 2007b. Linguistic and social predictors of copula use in Galician Spanish. *Spanish in contact: policy, social and linguistic inquiries*, ed. by Kim Potowski and Richard Cameron. Amsterdam: John Benjamins.
- 2008. Variation in contemporary Spanish: Linguistic predictors of estar in four cases of language contact. *Bilingualism: Language and Cognition* 11.365-80. Cambridge, MA: Cambridge University Press.
- Guijarro-Fuentes, Pedro, and Kimberly L. Geeslin. 2006. Copula Choice in the Spanish of Galicia. *Spanish in context* 3.63-83.
- Gutiérrez, Manuel J. 1992. The extension of estar: A linguistic change in progress in the Spanish of Morelia, Mexico. *Hispanic Linguistics* 5.109-41.
- 1994. Ser y estar en el habla de Michoacán, México. México, D.F.: Universidad Autónoma de México.
- Hopper, Paul and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language*. 56.251-299.
- Johnson, Daniel E. 2009. Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass*, 3(1):359-383.
- Luján, Marta. 1981. The Spanish copulas as aspectual indicators. *Lingua* 54.165-210.
- Malaver, Irania. 2000. "Cuando estábamos chiquitos": Ser y Estar en expresiones adjetivales de edad. Un fenómeno americano. *Boletín de Lingüística* 16.44-65.
- Mesa Alonso, M., Domínguez Herrera, M., Padrón Sánchez, E., and Morales Aguilera, N. 1993. Ser y estar. Consideraciones sobre su uso en español. *Islas* 104.150-56.
- Navas Ruiz, Ricardo. 1963. Ser y estar: estudio sobre el sistema atributivo del español. Salamanca: Acta Salmanticensia.
- Ortiz López, Luis A. 2000. Extensión de estar en contextos de ser en el español de Puerto Rico: ¿evaluación interna y/o contacto de lenguas? *BAPLE* 98-118.
- Pountain, Christopher. 1982. ESSERE/STARE as a Romance phenomenon. *Studies in the Romance Verb*, ed. by Nigel Vincent and Martin Harris, 139-60. London: Croom Helm.
- Resnick, Melvyn C. 1981. *Introducción a la historia de la lengua española*. Washington, D.C.: Georgetown University Press.
- Salazar, Michelle L. 2007. Está muy diferente a como era antes: Ser and Estar + adjective in New Mexico Spanish. *Spanish in contact: policy, social and linguistic inquiries*, ed. by Kim Potowski and Richard Cameron. Amsterdam: John Benjamins.
- Sanz, Cristina, and María José González. 1995. 'Ser' and 'estar' in Tortosi Catalan: Language Contact, Language Variation and Language Change. *Sintagma* 7.5.25.
- Scheibman, Joanne. 2001. Local patterns of subjectivity in person and verb type in American English conversation. In *Frequency and the emergence of linguistic structure*, ed. by Joan Bybee and Paul J. Hopper, 61-90. Amsterdam/Philadelphia: John Benjamins.

- Silva-Corvalán, Carmen. 1986. Bilingualism and language change: The extension of *estar* in Los Angeles Spanish. *Language* 62.587-608.
- . 1994. *Language Contact and Change*. Oxford: Clarendon Press.
- . 2001. *Sociolingüística y pragmática del español*. Washington, DC: Georgetown University Press.
- Torres Cacoullos, Rena & James Walker. In press. In *The Oxford Handbook of Grammaticalization*, ed. by Bernd Heine & Heiko Narrog. Oxford: Oxford University Press.
- Vañó-Cerdá, Antonio. 1982. *Ser y estar + adjetivos: un estudio sincrónico y diacrónico*. Germany: Gunter Narr Verlag Tübingen.
- Vergara Wilson, Damian. 2009. From 'remaining' to 'becoming' in Spanish: The role of prefabs in the development of the construction *quedar(se) + ADJECTIVE*. In *Formulaic language Vol. 1: Structure, Distribution, and Historical Change*, ed. by R. Corrigan, H. Ouali, E. Moravcsik, & K. M. Wheatley, 273-95. Amsterdam / Philadelphia: Benjamins.

# Selected Proceedings of the 14th Hispanic Linguistics Symposium

edited by Kimberly Geeslin  
and Manuel Díaz-Campos

Cascadilla Proceedings Project Somerville, MA 2012

## Copyright information

Selected Proceedings of the 14th Hispanic Linguistics Symposium  
© 2012 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-450-8 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.  
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

## Ordering information

Orders for the library binding edition are handled by Cascadilla Press.  
To place an order, go to [www.lingref.com](http://www.lingref.com) or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA  
phone: 1-617-776-2370, fax: 1-617-776-2271, [sales@cascadilla.com](mailto:sales@cascadilla.com)

## Web access and citation information

This entire proceedings can also be viewed on the web at [www.lingref.com](http://www.lingref.com). Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Brown, Esther L. 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*, ed. Kimberly Geeslin and Manuel Díaz-Campos, 61-74. Somerville, MA: Cascadilla Proceedings Project. [www.lingref.com](http://www.lingref.com), document #2656.