Copula Distinction and Constrained Variability of Copula Use in Iberian and Mexican Spanish

Sara Sánchez-Alonso  
_Yale University_

Ashwini Deo  
_Yale University_

María Mercedes Piñango  
_Yale University_

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/pwpl/vol23/iss1/25
For more information, please contact repository@pobox.upenn.edu.
Copula Distinction and Constrained Variability of Copula Use in Iberian and Mexican Spanish

Abstract
Spanish has two copulas, *ser* and *estar*, which are often translated as English ‘be’. Here, we study their differences by investigating their contrastive distributional patterns in combination with adjectival predicates. Specifically, we test the processing predictions of a presupposition-based analysis (Deo et al. 2016) that accounts for a wide range of distributional patterns of the copulas. This analysis has the advantage that it explains the variable copulas’ uses observed across Spanish varieties. Our focus is on Iberian and Mexican Spanish.

The presupposition-based analysis establishes a clear-cut distinction between the two copulas: *estar* presupposes the contingency of the prejacent, *ser* does not. Accordingly, the use of *estar* requires that the common ground contextually entails that its prejacent is contingent. If the common ground does not imply the contingency of the prejacent, this new information would need to be accommodated by the hearer. We hypothesize that *estar* predications, when presented in isolation with adjectival predicates that show a preference to appear with *ser*, will engender a processing cost as a result of adding to the common ground the proposition that the prejacent holds contingently.

This hypothesis is tested in two studies, an acceptability questionnaire and a self-paced reading. The results show that when the context does not explicitly support *estar*’s presupposition, sentences are scored lower (study 1, acceptability questionnaire) and read slower (study 2, self-paced reading) by both Iberian and Mexican speakers. In addition, the data provide experimental evidence for the ‘constrained’ variability across Spanish dialects. The results suggest that Mexican speakers are able to accommodate the contingency-presupposition of *estar* without relying on explicit contextual cues to a larger extent than Iberian speakers. Altogether, the data support an analysis of copula distinction in Spanish that takes into account the contingency-presupposition of *estar* and the variability in copula use across Spanish dialects.

This working paper is available in University of Pennsylvania Working Papers in Linguistics: http://repository.upenn.edu/pwpl/vol23/iss1/25
Copula Distinction and Constrained Variability of Copula Use in Iberian and Mexican Spanish

Sara Sánchez-Alonso, Ashwini Deo and María Mercedes Piñango

1 Introduction

Spanish has two copulas, ser and estar, which are often translated as English ‘be’. How do the range of semantic uses of ‘be’ get partitioned between the two Spanish copulas? We address this question by investigating the contrastive uses of the copulas and their variable acceptability across two Spanish varieties: Iberian and Mexican Spanish. The focus is on copula uses in combination with adjectival predicates. The use of ser has often been associated with predicating the existence of some property in the subject referent; whereas estar has been related to a variety of meanings (e.g., Ramsey 1894, Bull 1942, Crespo 1946, Roldán 1974a, Falk 1979, Franco and Steinmetz 1983). The examples below illustrate the variety of interpretations that estar has been associated with: it may denote 1) an accidental or temporary property, 2) a contrast with a prior known state of the referent, 3) a contrast with an expected state, or 4) the speaker’s subjective evaluation.

(1) a. La manzana es/está ser/estar PRES.3SING verde.
   The apple ser/estar PRES.3SING green
   ser: ‘The apple is green.’ (kind of apple)
   estar: ‘The apple is green.’ (it is not ripe yet)

b. ¡La carretera es/está ser/estar PRES.3SING ancha!
   The road ser/estar PRES.3SING wide
   ser: ‘The road is wide!’ (in comparison to the average width of roads)
   estar: ‘The road is wide!’ (in contrast to other parts of the same road)

c. Los zapatos (me) son/están ser/estar PRES.3PL pequeño.
   The shoes ser/estar PRES.3PL small
   ser: ‘The shoes are small.’ (kind of shoe)
   estar: ‘The shoes fit small.’ (relative to my foot size)

d. ¡La Capilla Sixtina es/está ser/estar PRES.3SING hermosa!
   The Sistine Chapel ser/estar PRES.3SING beautiful!
   ser: ‘The Sistine Chapel is beautiful.’ (objective statement)
   estar: ‘The Sistine Chapel is beautiful.’ (subjective statement)

Early grammarians noted that the most obvious distinction between the two copulas is that estar is typically used with predicates denoting temporary properties, whereas ser occurs with predicates that denote permanent properties (Ramsey 1894, Hanssen 1913, Roca Pons 1960, García de Diego 1970). This is the case in example (1a), in which the use of ser indicates that green is the color of the apple, (even) when it is in the fully ripe state; whereas estar is used to express that the apple is temporarily green as it transitions to becoming ripe. These first insights led researchers to analyze the copulas as lexical exponents of the stage-level versus individual-level distinction (Diesing 1992, Escandell-Vidal and Leonetti 2002). This dichotomy can account for a wide range of copula uses, but falls short in explaining several aspects of the distributional patterns, such as the examples (1b-d) above. More recently, some authors have argued that the differences in the temporal implications of ser/estar predications arise from aspectual differences between the two copulas. Some of these analyses derive this from the lexical content of the copulas (Luján 1981, Fernández Leborans 1995, Roby 2009, Marín 2010, Gumiel-Molina and Pérez-Jiménez 2012); whereas others argue that copula differences are the result of the specific syntactic features and syntactic relationships associated with each predication (Gallego and Uriagereka 2009, Zagona 2012, Fábregas 2012, Camacho 2012).

A relevant observation is that the distributional patterns of the copulas seem to vary across Spanish varieties. The use of estar in (1c-d) above, for example, is more widespread in Mexican and
Venezuelan Spanish, but relatively limited in Iberian Spanish (Jonge 1992, Malaver 2009). In spite of these observations, the difference between ser and estar has often been investigated assuming a general pattern that holds in an idealized “Modern Spanish” (e.g., Falk 1979, Luján 1981, Clements 1988).

We believe there is much to be gained by further investigating the variable acceptability of ser/estar across Spanish dialects. Here, we test the processing predictions of an analysis of the ser/estar distinction formulated in terms of contextual conditioning in use of semantic variants. We test the predictions in Iberian and Mexican Spanish, two varieties that have been argued to differ in their copula uses. We use two different methodologies to explore the offline comprehension and real-time processing differences of the copulas: 1) an acceptability questionnaire and 2) a self-paced reading study. The paper is organized as follows. First, we describe a presupposition-based account of the Spanish copulas and explain how it accounts for the distributional patterns in (1a-d). Then, we present the processing predictions of the analysis, and the two studies that investigate these predictions. A discussion of the implications of these findings concludes the paper.

2 A Presupposition-Based Analysis of ser and estar

Deo and colleagues (2016) present a semantic analysis of copula distinction in Spanish that is based on the observation from previous literature that the use of estar conveys a strong link with a so-called specific situation (Roldán 1974a,b, Franco and Steinmetz 1983, Clements 1988, Fernández-Leborans 1995, Escandell-Vidal and Leonetti 2002, Maienborn 2005, Arche 2006, Marín 2010, Fábregas 2012, Gumiel-Molina, Silvia, Moreno-Quiben, Norberto, Perez-Jimenez 2015). Their proposal of the ser/estar distinction has the advantage that it unifies both the typical temporary/accidental interpretations associated with estar (1a) and the non-temporary interpretations (1b-d).

The specificity condition of estar has been characterized in different ways in the literature, usually in terms of a presuppositional component present in the lexical meaning of estar. A recent account, presented by Maienborn (2005), builds on Clement’s (1988) intuitions and proposes that the two copulas differ only with respect to the presuppositional content of estar. Estar lexicalizes a specificity presupposition, which restricts the predication to a particular discourse situation. The use of estar thus requires the existence of alternative discourse situations at which the predication does not hold. Crucially, ser remains neutral on this issue. The presence of the presupposition has a pragmatic effect: estar sentences often give rise to quasi-exhaustive inferences in context – as restricted claims that may not hold in temporally, spatially or epistemically different discourse situations. Alternative situations may be defined along at least three dimensions: temporal (2a), spatial (2b) or epistemic (2c):

(2) a. La manzana está verde.
   The apple estar.PRES.3SING green
   ‘The apple is green.’

b. La carretera está ancha.
   The road estar.PRES.3SING wide
   ‘The road is wide.’

c. Las hojas están amarillas.
   The leaves estar.PRES.3SING yellow
   ‘The leaves are yellow.’

(Maienborn 2005:172)

Sentence (2a) would be uttered in a situation in which the apple is green because it has not ripened yet. The contrast, in this case, is established with a later state of the apple in which the predicate does not apply to the subject referent. Sentence (2b) would be uttered in a situation in which the speaker has driven through the Panamericana road, starting in Argentina, and s/he is now near Lima. The use of estar serves to indicate that there might be other parts of the road that are not as wide (for example, near Buenos Aires). Finally, Maienborn takes example (2c) from (Querido...
1976), who argues that *estar* can be used in situations in which the speaker does not have enough information to determine whether the property applies intrinsically to the subject referent or not. Maienborn calls this use of *estar* the ‘discovery interpretation’; for instance, (2c) could be uttered by someone who travels to the Amazon and discovers a new species of a tree.

Maienborn’s analysis leaves certain questions unanswered, such as which properties of a specific discourse situation are crucial to identify a discourse situation and how they relate to the present state. Subjective evaluations, as in (1d) above, are also not directly addressed in her analysis.

Seeking to circumvent these limitations Deo et al. (2016) propose an analysis that makes explicit the situation restrictedness of *estar* predications. The truth value of *estar* predications is directly compared across possible indices of evaluation. A possible index $i$ is defined as a tuple $\langle t, w, l, c \rangle$, where $t$ is an interval, $w$ is a world, $l$ is a location and $c$ is a contextual-standard function that assigns to every gradable predicate $P$ a standard that determines the positive extension of $P$. The prejacent to *ser* and *estar* is a proposition – a function from indices of evaluation to truth values (type $\langle s, t \rangle$). What uniﬁes the observed uses of *estar* is that each of these uses exhibits sensitivity to some parameter of the evaluation index (the specific discourse situation). The *ser/estar* puzzle is thus solved by formally modeling this parametrized sensitivity for *estar*.

On this analysis then, both *ser* and *estar* combine with a property denoting expression $P$ and an individual denoting argument $x$ and assert that $[x]^{i} \subseteq [P]^{i}$. *Estar* differs from *ser* in presupposing that $P(x)$ holds contingently at an index $i$. *Ser* by contrast remains neutral on this. A proposition of the form $P(x)$ holds contingently at an index $i$ whenever there is an $i'$ accessible from the evaluation index $i$, varying only along one contextually-determined parameter, such that $[P(x)]^{i'} = 0$ (the proposition is false at $i'$). Such minimally different pairs in the accessibility relation are written as $R_{p}(i, i')$. The possible relations (*time, location* etc.) are functions that apply to an index and return the value for the relevant parameter for that index. Of relevance for our purposes are the following accessibility relations:

1. $\forall i, i': R_{t}(i, i') \leftrightarrow \text{time}(i) \supset \text{time}(i')$

   An index $i'$ ($\langle t', w, l, c \rangle$) is temporally accessible from $i$ ($\langle t, w, l, c \rangle$) iff the temporal interval $t'$ of $i'$ abuts the temporal interval $t$ of $i$ ($t'$ is immediately before or after $t$).

2. $\forall i, i': R_{l}(i, i') \leftrightarrow \text{location}(i) \supset \text{location}(i')$

   An index $i'$ ($\langle t, w', l, c \rangle$) is spatially accessible from $i$ ($\langle t, w, l, c \rangle$) iff the location $l$ of $i'$ abuts the location $l'$ of $i$. ($l'$ is spatially adjacent to $l$.)

3. $\forall i, i': R_{w}(i, i') \leftrightarrow \text{world}(i') \in \text{Sim}(\text{world}(i))$

   An index $i'$ ($\langle t, w', l, c \rangle$) is modally accessible from $i$ ($\langle t, w, l, c \rangle$) iff the world parameter $w'$ of $i'$ is among the most similar worlds to the world $w$ of $i$.

The lexical entries of the copulas are as in (6) and (7), in which the only difference is that *estar* presupposes that the embedded proposition holds contingently, while *ser* remains neutral on this issue.

\[ [estar] = \lambda P, x, e, \lambda l : \exists i'[R_{p}(i, i') \wedge [P(x)]^{i'} = 0], [P(x)]^{i} = 1 \]

\[ [ser] = P, x, e, \lambda l : P(x), [P(x)]^{i} = 1 \]

2.1 Application of the Contingency-Presupposition Analysis to *estar* Predications

The presupposition-based analysis is able to account for common interpretations of *estar* such as the temporariness interpretation (8a) and contrastive uses with prior known states of the referent (8b). It also explains more problematic cases, such as the use of *estar* to signal contrast with an expected state (8c) and to convey the speaker’s subjective evaluations (8d).

8. a. La manzana es/está verde.
   the apple *ser/estar.*PRES.3SING green
ser: ‘The apple is green.’ (kind of apple)
estar: ‘The apple is green.’ (it is not ripe yet)
b. ¡La carretera es/está ancha!
The road ser/estar.pres.3sng wide
ser: ‘The road is wide!’ (in comparison to the average width of roads)
estar: ‘The road is wide!’ (in contrast to other parts of the road)
c. Los zapatos (me) son/están pequeños.
The shoes cl.io ser/estar.pres.3pl small
ser: ‘The shoes are small.’ (kind of shoe)
estar: ‘The shoes are small.’ (relative to my show size)
d. ¡La Capilla es/está hermosa!
The Sistine Chapel ser/estar.pres.3sng beautiful!
ser: ‘The Sistine Chapel is beautiful.’ (objective statement)
estar: ‘The Sistine Chapel is beautiful.’ (subjective statement)

Estar is used with stage-level predicates, as in (8a), because for any index \( i \) at which the relevant predication holds, there are temporally accessible indices \( i' \) such that the predication is false at \( i' \). In this particular case, one can identify future temporal intervals at which the apple will not be green, which license the use of estar. In (8b), the use of estar is licensed by the presence of a spatially accessible index \( i' \) at which the predication is false. What allows access to such an index is the speaker’s knowledge of other spatial parts of the road in which the property of being wide does not apply to the same degree as in the evaluation index. Finally, examples (8c-d) with estar are acceptable because there are modally accessible indices at which the predication is false. Access to such indices comes from the doxastic alternatives of the speaker at the time before the utterance time. In (8c), there is an accessible index \( i' \) at which the shoes are bigger in comparison to the shoe size at the index of evaluation \( i \). In (8d), the doxastic alternatives of the speaker make accessible an index \( i' \) at which the Sistine Chapel is beautiful to a lower degree in comparison to the predication at \( i \). Such minimally different indices are identical to the evaluation index except for its world parameter \( w' \), which would be an element of the pre-experiential doxastic alternatives of the speaker.

3 Processing Predictions

The presupposition-based analysis establishes a clear-cut distinction between the two copulas: estar presupposes the contingency of the prejacent, ser does not. Accordingly, the use of estar requires that the common ground contextually entails that its prejacent is contingent. If the common ground does not already imply the contingency of the prejacent, this new information would need to be accommodated by the hearer. Independently, it has been shown that presupposition accommodation is potentially measurable in behavioral studies, such as acceptability ratings, as well as in online processing, such as reading times and eye-tracking (see Schwarz 2016 for an overview). We hypothesize that the cost of accommodation in estar predications, when presented in isolation, results from adding to the common ground the proposition that the prejacent holds contingently along a contextually determined parameter of evaluation. Specifically, we test this hypothesis in cases in which estar occurs in sentences containing predicates that when presented without any contextual support, have a preference to occur with ser.

For our acceptability task (study 1), the prediction is that specific contextual support should result in higher ratings for estar predications when the sentence is preceded by a context that establishes the contingency of the prejacent. Such a context provides information about alternative indices at which the prejacent is false, thus explicitly updating the common ground to satisfy estar’s presupposition. By contrast, neutral contexts that do not make any claim regarding the contingency of estar’s prejacent should yield lower ratings.
For our self-paced reading study (study 2), we expect the cost of presupposition accommodation to translate into higher reading times when the context does not explicitly support estar’s presupposition in contrast to when an estar predication is preceded by a supporting context.

In both cases, our main prediction for ser is that no differences should be observed regardless of the context because ser is neutral with respect to the contingency presupposition. However, given that both ser and estar are competing in the lexicon for the same semantic ground and estar is predicted to have a more specialized meaning, we might also expect speakers to show a preference for ser to be used with neutral contexts.

4 Study 1. Acceptability Questionnaire

4.1 Materials and Design

Sentence pairs were created using Iberian Spanish (we will call each sentence context and test sentence, respectively). Sentences were normed for acceptability with four native speakers of Iberian Spanish naive to the purposes of the experiment. The same sentences were adapted to the variety of Mexican Spanish and normed by four native speakers.

The test sentences consisted of adjectival predicates with either ser or estar and were all in the present tense. All adjectives included can appear with both ser and estar, but show a preference for ser when presented in isolation. Every test sentence was preceded by a context sentence. Context sentences were of two kinds: they either made accessible indices that minimally differ from the evaluation index and where the predication was understood to be false (supporting context [9a]) or they did not imply the falsity of the predication at minimally different indices, thus being neutral with respect to the contingency presupposition (neutral context [9b]). Neither ser nor estar appeared in the context sentences to avoid biasing the participant towards a particular copula use.

(9) a. Supporting context: Durante el invierno, la arena blanca de la playa de Nogales, en Canarias, se vuelve de color gris.
   ‘During the winter, the white sand in Nogales beach, in the Canary Islands, turns grey.’

b. Neutral context: La playa de Nogales en Canarias, con su arena gris, se ha convertido en un gran atractivo turístico en cualquier época del año.
   ‘Nogales Beach, in the Canary Islands, with its grey sand, has become a tourist attraction throughout the year.’

Test sentence: Tienes que verla, la arena es/está gris por la acción volcánica de la isla.
   You’ve got to see it, the sand ser/estar grey because of the volcanic activity on the island.’

Thirty test sentences were constructed with ser and 30 with estar. Each test sentence was presented twice: once with a supporting and once with a neutral context. This set-up resulted in a 2x2 design with copula type as first factor and context type as second factor, for a total of 120 sentence pairs.

The Iberian Spanish version included 80 filler sentences that followed the same design as the experimental items, with the exception that the copula was followed by a noun or prepositional phrase. The Mexican Spanish version included, additionally, 240 fillers from an unrelated experiment. This resulted in a final script of 440 items for this variety (120 experimental sentences + 80 filler counterparts + 240 additional fillers). To ensure that participants were paying attention, a subset of the sentences was followed by a comprehension question. The Iberian version had 50 questions, 25% of the total number of items; the Mexican version had 220 questions, which constituted 50% of all items in the study.

4.2 Participants

Iberian Spanish: Forty monolingual speakers of Iberian Spanish (24 women) participated in the
experiment. The participants were between 19-36 years old (average age: 29) and had at least a secondary education. They were born in Spain and had lived in Spain most of their lives.

**Mexican Spanish:** Forty monolingual speakers of Mexican Spanish (22 women) participated in the experiment. The participants were between 19-37 years old (average age: 28) and had at least a secondary education. They were born and lived in Mexico City or surrounding states (Morelos or Puebla) and had lived in Mexico most of their lives.

### 4.3 Procedure

Participants read pairs of sentences (a context sentence followed by a test sentence with either *ser* or *estar*) and had to assign a number from 1 (non-native like) to 5 (native-like) according to whether or not they thought the sentence could be said by a native speaker of their own Spanish variety. They were asked to answer a series of comprehension questions after some of the sentences. Participants only saw sentences from their own variety of Spanish.

### 4.4 Results

We used R (R Core Team, 2012) and lme4 (Bates, Maechler and Bolker, 2012) to perform a multilevel analysis of the relationship between copula and context acceptability. As fixed effects, we entered copula, context and the interaction between context and copula into the model. As random effects, we had intercepts for subjects and items, as well as by-subject and by-item random slopes for the effect of copula and context. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question.

These are our findings: The context by copula interaction significantly predict acceptability scores across the two varieties of Spanish (Iberian Spanish: $\chi^2(1)=274.12$, $p<.001$; Mexican Spanish: $\chi^2(1)=34.81$, $p<.001$) This interaction was broken down by conducting separate multilevel models on ‘ser’ and ‘estar’. The models specified were the same as the main model, but excluded the main effect and interaction term involving copula type. The analyses showed that context type significantly predicted acceptability scores for each copula type: acceptability scores were higher for *estar* when the sentence was preceded by a supporting context, (Iberian Spanish: $\chi^2(1)=25.32$, $p<.001$; Mexican Spanish: $\chi^2(1)=4.8$, $p=.02842$); acceptability scores were higher for *ser* when the sentence was preceded by a neutral context, (Iberian Spanish: $\chi^2(1)=14.42$, $p=.0001432$; Mexican Spanish: $\chi^2(1)=5.5$, $p=.01894$). Mean scores for each variety are shown in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Iberian</th>
<th>Mexican</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supporting+Estar</td>
<td>3.74 (1.28)</td>
<td>4.19 (.97)</td>
</tr>
<tr>
<td>Neutral+Estar</td>
<td>3.04 (1.27)</td>
<td>3.92 (1.11)</td>
</tr>
<tr>
<td>Neutral+Ser</td>
<td>4.35 (.92)</td>
<td>4.44 (.77)</td>
</tr>
<tr>
<td>Supporting+Ser</td>
<td>3.7 (1.22)</td>
<td>4.24 (.93)</td>
</tr>
</tbody>
</table>

Table 1: Mean Acceptability Scores (SD) for each Spanish Variety by Sentence Type.

Adding Spanish variety as a fixed effect to the main model resulted in significantly different mean scores for *estar* sentences across the two varieties for both context types: Mexican variety > Iberian variety (Supporting context: $\chi^2(1)=16.42$, $p<.001$; Neutral context: $\chi^2(1)=41.26$, $p<.001$). With respect to *ser* sentences, no differences were found between the two varieties ($\chi^2(1)=1.72$, $p=.19$) when the sentences were preceded by neutral contexts. If *ser* combined with presupposition-supporting contexts, Mexican average scores were significantly higher than those provided by Iberian speakers ($\chi^2(1)=20.18$, $p<.001$): Mexican variety > Iberian variety. Figure (1) shows mean scores for each sentence pair combination by variety.
4.5 Discussion

The questionnaire data show that there is contextual modulation in copula acceptability across Spanish varieties: 1) 
estar-predications are rated higher when preceded by contexts that explicitly support its presuppositional component and 2) ser-predications receive higher scores when preceded by neutral contexts. In addition, we found variability in the rate of acceptability of estar predications between Mexican and Iberian Spanish speakers. Specifically, Mexican speakers assign higher ratings to estar predications regardless of context type.

This data provide experimental support for the observation, already made in the literature, that the distribution of the two copulas vary across Spanish dialects. Crucially, the results suggest that the variability is constrained. Ratings for estar predications are higher when the context explicitly provides information about contrasting alternatives that allow the predication to be construed as contingent. This contextual modulation of acceptability ratings serves as support for an analysis of the ser/estar distinction that takes into account the contingency-presupposition of estar.

5 Study 2. Self-Paced Reading

5.1 Materials and Design

The materials were the same as in Study 1. The number of items was increased to 180 sentences.

5.2 Participants

The criteria for participant’s recruitment were as in Study 1. A total of 101 participants were tested: 61 speakers of Iberian Spanish (33 women) and 40 speakers of Mexican Spanish (29 women).

5.3 Procedure

Sentences were presented word by word using the moving-window technique and reading times were recorded for each displayed segment. The critical regions were defined as the copula, the adjective immediately following the copulas, and the next two subsequent regions. After each sentence, a yes/no question or statement about the sentence was presented to ensure that participants paid attention to the sentences and processed them fully. Apart from the experimental items, there were 90 items from an unrelated experiment in the Iberian version of the questionnaire and 144 items, also from an unrelated experiment, in the Mexican version.

5.4 Results

Reading times greater than two standard deviations below or above the mean for each condition and for the four critical segments were excluded from the analysis (17% of the data). The mean
percentage of response accuracy among participants was 90%.

For the Iberian data, a multilevel analysis on context and copula type showed an interaction of the two predictors at two words after participants encountered the adjective ($\chi^2(1)=10.39, p<.001$). The interaction was broken down by conducting separate multilevel analyses on ser and estar, which showed that estar sentences were read slower when preceded by a neutral context ($\chi^2(1)=15.02, p=.001$), whereas no differences were observed between ser sentences ($\chi^2(1)=.5, p=.47$). No other significant interactions were found at any of the other sentence segments (COP: $\chi^2(1)=.66, p=.42$, ADJ: $\chi^2(1)=.7, p=.4$, ADJ+1: $\chi^2(1)=.77, p=.38$). The Mexican data did not show any significant differences across sentence pairs for any of the crucial segments.

![Figure 2: Mean Reading Times for each Context-Copula Pair by Segment (Iberian Spanish).](image)

![Figure 3: Mean Reading Times for each Context-Copula Pair by Segment (Mexican Spanish).](image)

### 5.5 Discussion

Results for Iberian Spanish indicates that estar sentences engender higher cost when preceded by contexts that do not provide explicit cues to contrasting alternative indices. Estar predications are read at a similar rate as ser sentences if they are preceded by a contingency-supporting context. By contrast, results for Mexican speakers did not show any differences in reading times for ser/estar predications, regardless of context type. From these results, we can conclude that Iberian speakers tend to rely more on explicit presupposition-supporting cues during processing of estar predications than Mexican speakers.

### 6 General Discussion

We have investigated the processing predictions of a presupposition-based analysis (Deo et al., 2016) of the ser/estar distinction in Iberian and Mexican Spanish. We capitalize on two different methodologies: an acceptability questionnaire and a self-paced reading study. The presupposition-based
account establishes a clear-cut distinction between the two copulas: *estar* encodes a contingency presupposition, whereas *ser* is neutral on this respect. This account has the advantage of being able to explain a wide range of distributional patterns and the variable copula uses observed across Spanish varieties, in addition to making processing predictions.

The results from the questionnaire indicate that *estar* predications receive lower scores when they are preceded by neutral contexts that do not provide explicit cues to the contingency of the prejacent. We interpret this as an effect of accommodating the contingency-presupposition of *estar*: if the context does not explicitly support the contingency of the prejacent, speakers need to add this information into the common ground, which leads to lower acceptability ratings. Although contextual modulation for *estar* sentences was found in the two Spanish varieties, Mexican speakers tend to assign higher scores to *estar* sentences regardless of context type.

The data from the self-paced reading study, which provide a measure of copula processing online, as the sentence unfolds, shed further light on copula differences and their variable patterns across the two Spanish dialects. We found that only Iberian speakers show a processing cost for *estar* predications when preceded by a neutral context, in comparison to *ser* sentences and *estar* predications preceded by contingency-supporting contexts. This data align with the results from the questionnaire, thus providing support for the additional processing cost of *estar* sentences when the information regarding the contingency of the prejacent needs to be added to the common ground. By contrast, Mexican speakers do not show any processing cost for *estar* sentences regardless of context type. We argue that the Mexican variety is less dependent on explicit contextual cues on the contingency of *estar*’s prejacent.

Differences between *ser* sentences were only found in the questionnaire data. In both Iberian and Mexican Spanish, speakers tend to provide higher ratings to *ser* predications that are preceded by a neutral context. We argue that the contextual modulation of *ser* sentences is the result of the division of labor between the two copulas: *estar* predications are used to express contingent propositions, thus becoming the preferred copula with contingency-supporting contexts.

Altogether, the experimental data support the analysis that *estar* predications presuppose the contingency of the prejacent. As such, speakers benefit when information about contrasting alternative indices at which the prejacent is false is explicitly provided by the context. The lack of explicit contextual cues to the possibility of contingency translates into lower acceptability scores and higher reading times. In addition, the studies provide experimental support, both online and offline, for the constrained variability in copula use for Iberian and Mexican Spanish. The data suggest that Mexican speakers are able to accommodate the contingency-presupposition without relying on explicit contextual cues to a larger extent than Iberian speakers. All in all, this data support an analysis of the *ser*/*estar* distinction that takes into account the presuppositional content of *estar* and is able to explain the variable copula uses across Spanish dialects.

References


Roldán, María M. 1974b. Toward a semantic characterization of Ser and Estar. Hispania 68–75.
