Ser and Estar in Spanish: A Scalar Account

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Carmelo Bazaco, MA
Graduate program in Spanish and Portuguese

The Ohio State University

2017

Dissertation Committee:

Rebeka Campos-Astorkiza, Adviser
Peter W Culicover
Melvin González-Rivera
Kiwako Ito
Fernando Martínez-Gil
Abstract

The distribution of the copulas *ser* and *estar* presents an interesting problem in Spanish. Although there is a vast literature on the topic, the difference in their interpretations remains elusive. This dissertation proposes that all uses of *ser* and *estar* can be reduced to a difference in the scalar properties of their predicates.

Several problems are identified in the existing literature regarding copula distribution in Spanish. Traditional approaches to *ser* and *estar* base their accounts on temporal properties, instantiated around concepts such as permanent and temporary, stage and individual level predicates or other aspectual properties. A close analysis of *ser* and *estar* predicates reveals that temporal properties are not a factor in their distribution. A sentence such as *Esas casas rojas están vacías* ‘Those red houses are empty’ gives no information about the occupancy of the houses at any point in time. Another area where traditional theories also make wrong predictions about the distribution of the copulas. Adjectives like *rico* ‘rich’ and *famoso* ‘famous’ can easily be interpreted as both permanent or temporary, and yet their use in *estar* predications is unobserved in Spanish. Any theoretical approach to *ser* and *estar* needs to be able to explain these gaps in the data. Finally, evidential uses of *estar*, such as *Este vino está estupendo* ‘This wine is great’ not only escape traditional accounts, but have mostly been ignored, despite being commonplace in Spanish.

This dissertation proposes that *ser* predicates denote a single point on a scale, whereas *estar* predicates denote an interval on that scale. In the case of non-evidential uses of *estar*, I propose that *estar* is sensitive to the scalar structure of adjectives. In order to generate a scalar interval, closed-scale adjectives, such as *vacio* ‘empty’, require access to the opposite state defined by the adjective. Open-scale adjectives, like *guapo* ‘pretty’ also require a natural process that drives the change in state. The importance of this natural process is not only shown based on data from Spanish, but also from the empirical data gathered in the experimental section of this dissertation. In evidential *estar*, the scalar interval is generated by comparing the distance of the degree projected by the subject of the predication to the standard of comparison.

The theoretical framework in this dissertation accounts for the non-temporality of the copulas by requiring *estar* predications access to a potential different state, not an actual realization of it. Temporary predicates become a sub-set of all uses of *estar*. Additionally, it can account for the gaps in the data in terms of the requirement of a natural process in open-scale *estar* predications. Finally, evidential uses of *estar* are subsumed under the same paradigm as non-evidential uses, the only difference being how the scalar interval is constructed. Although this dissertation focuses on adjectival predications, its proposal can be extended to other contexts of *ser* and *estar*. Furthermore, the relevance of the natural process and the result of the experiment presented here presents future avenues for research in L2 acquisition and pedagogy.
Vita

June 1994 ........................................................................................................Lincoln County High School

2000........................................................................................................B.A. English Philology, University of Deusto

2010...........................................................................................................M.A. TSL Spanish, Fundación Comillas

2012.............................................................................................................M.A. Hispanic Linguistics. The Ohio State University

2012 to present........Graduate Teaching Associate, Spanish and Portuguese Department, The Ohio State University

Publications


Fields of Study

Major Field: Spanish and Portuguese
Table of Contents

Abstract...............................................................................................................................................i
Vita....................................................................................................................................................ii
List of Tables......................................................................................................................................iv
List of Figures.....................................................................................................................................v
Chapter 1. Distribution of ser and estar............................................................................................1
Chapter 2. Previous Theoretical Accounts..........................................................................................41
Chapter 3. Scalar Account for Ser and Estar......................................................................................86
Chapter 4. Online Processing of Ser and Estar..................................................................................173
Chapter 5. Conclusions.....................................................................................................................220
References.........................................................................................................................................239
Appendix A. Target Items for Experiment.........................................................................................251
List of Tables

Table 1. Distribution of *ser* and *estar* ......................................................................................................................... 40
Table 2. Experiment 1 results from Holtheuer et al. (2011). In the Ø option, no answer was given .................. 182
Table 3. Table 3. Experiment 2 results from Holtheuer et al. (2011) .............................................................. 183
Table 4. Distribution of *estar* + color adjectives across the CREA and Davis corpora ............................... 189
Table 5. Background information of all participants ................................................................................................. 191
Table 6. Background information of participants who met the clicking criteria .................................................. 192
Table 7. Summary of mixed effect regressions on the fixation ratios. All participants ........................................ 204
Table 8. Background Information of Participants per Subgroup ........................................................................... 207
Table 9. Summary of mixed-effect regressions on the fixation ratios ............................................................. 211
Table 10. Summary of pairwise analysis on the fixation ratio ................................................................................. 212
Table 11. Summary of correlation between log ratio looks to the natural picture and the distractor and the accuracy in the clicking task data .......................................................................................... 214
Table 12. Clicking Accuracy ........................................................................................................................................... 215
Table 13. Experiment 1 results from Holtheuer et al. (2011) ............................................................................... 216
Table 14. Accuracy Task Errors per Condition ............................................................................................................. 216
List of Figures

Figure 1. Sample target pictures and story from Schmitt & Miller (2007)..................................................177
Figure 2. Results Experiment 2 from Schmitt & Miller (2007). Percentage of copula used per condition........178
Figure 3. Same character and different character condition. From Holtheuer et al (2011)............................181
Figure 4. Top left corner is the estar + natural process condition (dirty fingernail), top right is the estar + artificial process condition (painted fingernails), bottom left is the ser condition (gorilla hands) and bottom right is the distractor...............................................................196
Figure 5. Fixation rates for the ser condition..............................................................................................200
Figure 6. Fixation rates for the estar with natural process condition.............................................................201
Figure 7. Fixation rates for the estar with artificial process condition..........................................................201
Figure 8. Window 1 Fixations...................................................................................................................202
Figure 9. Window 2 Fixations...................................................................................................................203
Figure 10. Window 3 Fixations..................................................................................................................203
Figure 11. Histogram of participants according to accuracy........................................................................207
Figure 12. Higher Accuracy Group (Left), Middle Accuracy (middle) and Loer Accuracy Group (right). Ser Condition.................................................................................................................................209
Figure 13. Higher Accuracy Group (Left), Middle Accuracy (middle) and Loer Accuracy Group (right). Estar with Natural Process Condition........................................................................................................210
Figure 14. Higher Accuracy Group (Left), Middle Accuracy (middle) and Loer Accuracy Group (right). Estar with Artificial Process Condition........................................................................................................210
Figure 15. Accuracy in Clicking Task and log ratio between looks to natural process picture and distractor....213
Chapter 1: Distribution of *ser* and *estar*

1.1 Introduction

Every human language has, in one way or another, a way to establish a connection between an entity, be it animate or inanimate, and the qualities it possesses or the state it finds itself in. Many languages do so by means of a copular verb, for example the verb *to be* in English. In the case of Romance languages, we can find languages that only have one copula—*être* in French, and *fire* in Romanian—and languages that have two—*essere* and *stare* in Italian, *siri* and *stare* in Sicilian. Spanish belongs to the second group, and has a dual copula system, *ser* and *estar*.

The two copular verbs in Spanish—*ser* ‘be*SER*’ and *estar* ‘be*ESTAR*’—have long attracted the interest of scholars. Their distributional properties are still one of the most studied features of Spanish grammar, as evidenced by the multitude of authors and theoretical approaches that the topic has received—Luján (1981), Diesing (1990, 1992), Kratzer (1989, 1995) from a semantic/aspectual perspective; Gumiel et al (2015) from a scalar one; Maienborn (2003, 2005) pragmatic; Zagona (2009, 2010), Gallego & Uriagereka (2009, 2016), Camacho (2012) syntactic, to name a few. It is both a central part of the Spanish language, both copulas being extremely frequent in use, and yet one of the hardest to completely subsume under one single framework, as anyone who has had to learn Spanish can attest. Although there is a vast literature on the topic—ranging from the merely descriptive accounts and the historical evolution of both copulas, to their dialectal variation and their acquisition by L2 learners of Spanish—there is no account that can be said to capture the exact conditions for the distribution of *ser* and *estar* in a way that
generally satisfies linguistic theorists. Every single of the proposed theoretical treatments of the
distribution of *ser* and *estar* fails in explaining or predicting the behavior of certain structures
found in the language. And, frustratingly for many L2 learners, many of those unaccounted-for
structures are indeed quite common in Spanish.

From a theoretical point of view, the existence of this dual copula system in Spanish provides a
variety of contexts, as Roby (2009) points out, where several relevant linguistic issues can be
observed. For example, it can illuminate the interaction between syntactic, semantic and
pragmatic factors which explains the properties of copular sentences, given that both copulas, *ser*
and *estar*, appear in virtually any type of predication in Spanish. Also, the diachronic
development of the copular system can shed some light in language evolution at all linguistic
levels. Furthermore, studying the acquisition and processing of copular sentences, both in L1 and
L2 contexts, can be very fruitful in terms of language acquisition and second language teaching.

It is precisely in this last field where a concise, all-encompassing theoretical account of the
distribution of both copulas can play a central role outside the world of theoretical linguistics. As
previously mentioned, explaining when to use *ser* and when to use *estar* is one of the main
challenges that both instructors and students face in a Spanish classroom.

The main objective of the present dissertation is to develop a theoretical account that acts as a
reliable predictive tool for what structures will and won’t be found with *ser* and *estar* in Spanish.
As will be pointed out in later sections, a recurrent problem with some of the theoretical accounts
available is that they explain the existing data while also predicting unobserved structures. This

2
dissertatio will also attempt to incorporate all the relevant contributions that the main existing accounts provide. In this sense, it brings under one single theoretical umbrella the distilled essence of previous accounts. Finally, all of these efforts will be guided by a principle of economy. The present account will try to reduce the distribution of *ser* and *estar* to just one linguistic mechanism, and to dress this mechanism in as little artifice as possible.

In Chapter 1, I provide an inventory of all the contexts where *ser* and *estar* can be found in Spanish, together with a brief explanation in terms of the classical permanent vs temporary distinction. The objective of this first chapter is to provide a baseline of knowledge about the copulas in Spanish, paying special attention to the contexts in which they can occur. Additionally, each context is also explained in the traditional terms of the permanent vs. temporary distinction for the copulas. Furthermore, this chapter also provides counterexamples to this traditional view, in order to obtain not only a global perspective on how *ser* and *estar* are distributed in Spanish, but also the problems that accounting for this distribution in a thorough manner presents.

Chapter 1 provides a foundation on which to build on in Chapter 2. The second chapter of this dissertation focuses on the theoretical descriptions of the phenomenon that have been formulated in the literature. This second chapter goes beyond the classic permanent vs. temporary distinction and presents theoretical approaches to the problem from a variety of perspectives: semantic, syntactic and pragmatic. Additionally, Chapter 2 not only provides an insight into the strong points of the different proposals presented, but also the shortcomings of said descriptions. The
final objective of chapter is to, based on the insight gained from the literature on the subjects, determine the characteristics that a theoretical account of ser and estar in Spanish needs to have. These characteristics are summarized in three observations: i) the distribution of ser and estar is not based on strict temporal considerations, but rather on potential states of an individual; ii) the distribution of ser and estar is sensitive to the scalar structure of the predicate they combine with; and iii) evidential uses of estar need to be taken into consideration and incorporated into the theory. These observations provide a basis for the theoretical account presented in Chapter 3, which is the semantic foundation at the core of this dissertation. In this third chapter, the observations gathered in Chapter 2 are incorporated into a proposal for the distribution of ser and estar with adjectives in Spanish, based on the scalar properties of those adjectives. The main claim put forward in Chapter 3 is that ser combines with predicates that denote a single degree on a scale, whereas estar combines with adjectives that denote an interval on a scale. This basic premise is extended to account for the non-temporal nature of the copulas, their scalar sensitivity and the evidential uses of estar, thus subsuming the distribution of both copulas in all possible adjectival predications under one paradigm. One of the observations, previously unnoticed in the literature, introduced in Chapter 3 is that open-scale adjectives require not only a scalar interval in order to appear in estar predications, but also a natural process that drives that interval. Chapter 4 takes that observation and determines the effect that those processes have in the online interpretation of estar predications involving color adjectives in Spanish. By means of an eye-tracking experiment, the reactions of native speakers of Spanish when listening to sentences containing color expressions containing ser, estar with a natural process, and estar without a natural process are examined. The results show how native speakers do show an increased
number of looks towards pictures depicting natural processes upon hearing the copula *estar*, which can be interpreted as the natural process having been lexicalized and incorporated to the meaning of the copula. Chapter 5 presents the final conclusions for this dissertation and provides future areas of research, based on both the theoretical as well as the empirical conclusions presented throughout the dissertation.

### 1.2 The Traditional Account: Temporary vs Permanent

One of the most common approaches to the *ser* and *estar* distinction is the one that associates permanent characteristics with *ser* and temporary ones with *estar* (Bello, 1951; Gili Gaya, 1955; Luján, 1981; Delbecque, 1997). Under this point of view, the key factor governing copula choice in Spanish would be the relative duration of the predicate. In the case of predicates that combine with *ser*, this duration is permanent. Consider the in (1)\(^1\):

\[
\begin{align*}
(1) & \quad a. \quad \text{El Sol es una estrella.} \\
& \quad \text{The Sun is a star} \\
& \quad \text{‘The Sun is a star’} \\
& \quad b. \quad \text{El Empire State Building es un edificio muy alto.} \\
& \quad \text{The Empire State Building is a very tall building} \\
& \quad \text{‘The Empire State Building is a very tall building’} \\
& \quad c. \quad \text{Marta está feliz.}
\end{align*}
\]

\(^{1}\) This dissertation is focused on the universal, synchronic uses of *ser* and *estar* within the Spanish speaking world. All examples provided will be of structures that can be found in the entire dialectal continuum. Examples displaying dialectal uses of either copula will be indicated as such.
Marta \textit{is ESTAR happy}

‘Marta is happy’

d. La puerta del garaje \textit{está abierta}.

The door of the garage \textit{is open}.

‘The garage door is open’

The examples in (1a) and (1b) depict very stable properties of those individuals—it is hard to think how the Sun could cease to be a star, or how the Empire State Building would not be considered a tall building. The predicates in (1c) and (1d), on the other hand, have a more mutable nature—\textit{being happy} is a state that is very much a transitory one, and so is a door being open. These types of examples, where permanent properties are in a \textit{ser} predication, and temporary ones in an \textit{estar} one, are commonplace and easy to find in Spanish, and as such, it is not surprising that, at first glance, the difference in the interpretation of both copulas could be considered to center around the relationship between temporal nature of the predicate and its subject.

An extension to the permanent vs. temporary explanation is provided by Gil Gaya (1955), in another classic work. According to this author, \textit{ser} combines with inherent properties of individuals, whereas \textit{estar} does so with circumstantial ones. This theory is not without problems, however. It is attempting to explain the distribution of the two copulas based on a mostly perceptual, philosophical basis. Who or what determines the inherent qualities of an entity? It can be argued that certain material properties—clouds are made of water vapor, a tree
trunk is made of wood, water is wet and ice is cold—could be construed as essential to the individual. Nevertheless, there are many cases where Spanish demands the copula *ser* with predicates that are hard to argue as permanent to the individual (2a), and *estar* with predicates that are hard to be considered as temporary (2b):

(2)  

a. Marta es inteligente.  

Marta *is* *ser* intelligent  

‘Marta is intelligent’


Buenos Aires *is* *estar* in *Argentina*  

‘Buenos Aires is in Argentina’

Yet another attempt to extend the permanent vs temporary interpretation of *ser* and *estar* is considering *estar* as an indicator of a change of state (Bull, 1942), and *ser* an elsewhere option. This proposal has the benefit of accounting for some contexts of *estar* where a simply temporary consideration fails. For example, one of the prototypical counterexamples given against the temporary consideration of *estar* predicates is *muerto* ‘dead’. It is hard to argue that a dead person is going through a temporary condition, or that it is circumstantial, and yet it is a strict *estar* context. However, if *estar* is interpreted as involving a change of state, then (3) can be explained.
Under a strict *estar* for temporary predicates, being dead is permanent, so it should not work as an exclusively *estar* adjective, as it happens in Spanish. If, on the other hand, the meaning *estar* is understood to refer to predicates that have changed or can change, being dead is a consequence of no longer being alive. In that sense, this use of *estar* can be accounted for. Later in this chapter, we will see that this change of state consideration is not without problems, either.

### 1.3 Distribution of *ser* and *estar* in Spanish

In the following sections, the distribution of both copulas, *ser* and *estar*, will be examined in all the contexts in which they appear. In order to obtain a general perspective not only of where *ser* and *estar* appear, but also what some of the general problems that they present, each context also includes an analysis based on the traditional theoretical account. This will provide the dual benefit of having a contextual basis of the phenomenon, as well as a perspective on the problems that it presents.

After these contexts have been examined under the traditional lens, I provide a series of contexts that are problematic under this framework. These challenges include both classical counterexamples, as well as some that have received less attention in the literature.
1.3.1 *Ser* and *estar* with adjectives

Perhaps the context that has received the most attention in the literature has been the distribution of *ser* and *estar* with adjectival complements. Adjectives denote a wide range of qualities: temporal, locative, inherent and transient qualities, etc. In this manner, they provide a very fertile ground on which the *ser* and *estar* distinction can be made apparent. Furthermore, adjectival complementation (4) is the only context in which we can find the same complement that can potentially combine with both *ser* and *estar*. Although it is true that Noun Phrases (5) and Prepositional Phrases (6) can appear with both *ser* and *estar*, it is never the case that they can combine with either copula:

(4)  
\[ \text{a. Juan es} \{aburrido/alto/rojo\}. \]
\[ \text{Juan is} \text{SER}\ \{\text{bored/tall/red}\} \]
\[ \text{‘Juan is} \{\text{bored/tall/red}\}.’ \]
\[ \text{b. Juan está} \{aburrido/alto/rojo\}. \]
\[ \text{Juan is} \text{ESTAR}\ \{\text{bored/tall/red}\} \]
\[ \text{‘Juan is} \{\text{bored/tall/red}\}.’ \]

(5)  
\[ \text{a. Juan } \{\text{es/*está}} \text{ un estudiante.} \]
\[ \text{Juan } \{\text{isSER} / \text{isESTAR}\} \text{ a student} \]
\[ \text{‘Juan is a student.’} \]
These two features—range of qualities, and alternation between copulas—provide an ideal vantage point from which to observe the differences between *ser* and *estar*. It is therefore not surprising that most authors focus their attention in trying to determine the rules that govern the two copulas and their adjectival complements.

In general terms, adjectives in Spanish can be divided into three categories depending on their choice of copula: those that combine with *ser* exclusively (7a), those that combine with *estar* (7b) exclusively, and those that can combine with both (7c).

(7)  

a. Marta es {inteligente/guapa}.  
Marta is *SER {intelligent/pretty}

‘Marta is {intelligent/pretty}’

b. Marta está {enferma/borracha}.  
Marta is *ESTAR {sick/drunk}
‘Marta is {sick/drunk}’
c. Marta {es/está} {feliz/aburrida}
Marta isSER,iESTAR {happy/bored}
‘Marta is {happy/bored}’

These examples are all in keeping with the traditional permanent/temporary distinction. Of course, it is data like the one presented in (7c) that is extremely relevant in any approach to the distribution of ser and estar, as it presents predications in which both are valid alternatives.

Capturing the context in which either the ser or the estar option is chosen is critical. If a difference in the interpretation of sentences like (7c) depending on whether ser or estar is used can be captured, we will be closer to obtaining a characterization of the differences between the two copulas. Furthermore, contrasting the properties of those adjectives that can be combined with ser and estar with those that are exclusive to one can also shed light on their distributional differences.

The first group of adjectives, such as those found in (7a), are those that can only appear in ser predications. They are, as previously mentioned in this chapter, traditionally associated with inherent characteristics of individuals. In the examples in (7), it is safe to assume that, if Marta is intelligent, or pretty, she will continue to be so indefinitely. Accordingly, they are prototypical examples of ser predications. Prototypical groups of adjectives that exemplify this category of copula distribution in Spanish also include adjectives describing material composition (8a) or shape (8b), as they are stable, inherent properties of individuals:
The second group of adjectives to consider are those that can only appear with *estar*, like *asustado* ‘frightened’ or, *lleno* ‘full’. *Estar* predications in this type of context are very productive with adjectives depicting physical states (9a), or properties such as fullness or occupancy (9b), as these tend to be of a temporary nature—glasses are filled and emptied and rooms are occupied and vacated:

(9)  

a. Carola está embarazada.  
Carola is *estar* pregnant  
‘Carola is pregnant.’  

b. El estadio está lleno.  
The stadium is *estar* full  
‘The stadium is full’
being happy or bored can be taken, under the right circumstances, to be either a permanent or a temporary quality of an individual—permanent interpreted as being inherent, and temporary as contingent. When we predicate of Marta that she *está aburrida* ‘she is *ESTAR* bored’, the interpretation leads to a temporary one—Marta is bored because, for instance, it is a rainy Sunday afternoon and she is stuck at home with nothing to do. If, on the other hand, we say that Marta *es aburrida* ‘she is *SER* bored’, the implication is that Marta is generally a boring person, rainy afternoon or not.

One interesting phenomenon that Fernández Leborans (1999), among others, points out is that adjectives *a priori* belonging to the *ser* exclusive group can very often be coerced into *estar* predications. For example, *intelligent* ‘intelligent’ is usually considered to belong in the *ser* only group of adjectives. However, examples of *estar with inteligente* can be forced in certain contexts. The following example (10) is taken from a news source:

(10) Scioli estuvo inteligente en no responder ninguna de esas chicanas.

Scioli was *ESTAR* intelligent in no respond none of these provocations

‘Scioli was intelligent enough not to respond to any of those provocations’

http://www.politicargentina.com/notas/201511/9732-macri-estuvo-pedante-y-agresivo-quoed-claro-que-su-
propuesta-es-el-ajuste-y-la-devaluacion.html
In example (10), the subject is described as having a temporary burst or spike of intelligent behavior, and, as such, the predicate inteligente ‘intelligent’ fits in a traditional interpretation of estar predication. This type of predication, where an individual temporarily increases a particular characteristic, is very productive in Spanish with adjectives that describe character or behavior traits, as the examples in (11) show:

\[(11)\]
\[
\text{Ayer, Carlota estuvo } \{\text{honesta/atrevida/arrogante/encantadora}\}. \\
\text{Yesterday, Carlota was } \text{ESTAR } \{\text{honest/daring/arrogant/charming}\}
\]

‘Yesterday, Carlota was \{honest/daring/arrogant/charming\}’

Fernández Leborans (1999) proposes that the acceptability of this type of coercion depends on the temporal nature of the change indicating a fleeting, burst-like nature of the change of state. This intuition seems to be confirmed when we compare short periods of an increased feature (12a) with longer ones (12b):

\[(12)\]
\[
a. \text{Carlota estuvo encantadora durante toda la cena de ayer.} \\
\text{Carlota was } \text{ESTAR } \text{charming during all the dinner of yesterday}
\]

‘Carlota was char{}ming during the whole dinner yesterday’

\[
b. \text{Carlota estuvo encantadora durante sus cinco años en OSU.} \\
\text{Carlota was } \text{ESTAR } \text{charming during her five years in OSU}
\]

‘Carlota was charming during her five years at OSU’
Fernández Leborans (1999) argues that the example in (12b) seems to indicate that the temporary and permanent distinction is operative here. Once a predicate moves towards being considered a less inherent one—Carlota being charming for one evening as opposed to five years—the acceptability of the coercion decays.

Fernández Leborans (1999) also points out to yet another interesting fact, more precisely that adjectives that describe inherent biological properties of a species do not accept this kind of coercion. For example, *vertebrado* ‘vertebrate’ describes the property of an animal possessing a spinal column. This is, of course, as permanent as a property can be, and there is no logical way of interpreting it as a valid *estar* predicate in Spanish (13):

(13) #El perro estuvo vertebrado ayer.
    The dog *was* vertebrate yesterday
    ‘The dog was vertebrate yesterday’

1.3.2 *Ser* and *Estar* with Prepositional Complements

Both *ser* and *estar* are capable of taking prepositional phrases in Spanish. At first glance, it also appears that the permanent vs temporary distinction is operative in these contexts. For example, prepositional complements denoting the material composition, as in (14a) and (14b), or origin (14c) of the subject, can only combine with *ser*:
a. La falda de Marta es de seda.  
The skirt of Marta is ser of silk  
‘Marta’s skirt is made of silk’

b. La mesa de la cocina es de madera.  
The table of the kitchen is ser of wood  
‘The kitchen table is made of wood’

c. Mi amiga es de Colombia.  
My friend is ser of Colombia  
‘My friend is from Colombia’

This distribution fits perfectly with the traditional description of a permanent vs temporary distinction for ser and estar. Both the composition of an object and a person’s origin are stable, inherent properties and are not subject to change in any reasonable way.

In contexts where the predication involves a locative element, the situation becomes more complex, since temporal and physical location exhibit completely different behaviors. A locative predicate is possible when estar is followed by locational Prepositional Phrases (15a), whereas it is blocked with directional ones (15b):

(15)  
a. El Arco de Triunfo está en los Campos Elíseos.  
The Arc of Triumph is estar in the Champs Élysées  
‘The Arc of Triumph is in the Champs Élysées’
b. *El Arco de Triunfo está desde la Plaza de la Concordia.

The Arc of Triumph is from the Place of the Concorde

‘The Arc of Triumph is from the Place de la Concorde’

It is worth pointing out that in sentences such as (15b), *ser* is also unacceptable, which leads us to believe that *ser* and *estar* in Spanish are not it complimentary distribution in all cases.

Porroche (1990) proposes that Spanish considers all cases of locative complements as temporary in nature, whether our knowledge of the world agrees or not with this judgement. This consideration can be traced back to Latin and its use of *stare*—the etymological origin of *estar*—as the copula of choice for this type of construction. In cases of physical location, the nature of the subject, as Luján (1981) points out, can affect the choice of copula. If the location of the individual involved is mobile, we can find only *estar* as the copula. If the location is, on the other hand, fixed, *ser* becomes an option. The examples in (16) show this distribution:

(16) a. La oficina de correos {está/es} a la vuelta de la esquina.

The office of post {is/ESTAR/ISER} at the turn of the corner

‘The post office is around the corner’

b. Las tazas de café {están/*son} en el lavavajillas.

The cups of coffee {are/ESTAR/*areSER} in the dishwasher

‘The coffee mugs are in the dishwasher.’
In (16a), the subject is one that cannot logically change its location, as it is not plausible to think of a post office that moves around. In this case, both *ser* and *estar* are valid options, with no apparent difference in meaning, as Luján (1981) points out. It is also worth mentioning at this point that this use of *ser* in locative predications is less frequent than the *estar* alternative. What seems clear is that locative expressions of this kind do not seem to conform to the traditional distinctions. If, on the other hand, the subject of the predication is an entity that could conceivable change the location expressed in the predication, (16b), only *estar* is a valid option. In that particular context, the permanent and temporary distinction does seem to be the factor that drives, at least partially, the distribution of the copulas.

Temporal prepositional predications, on the other hand, are strictly *ser* contexts. This distribution can be seen in the examples in (17):

(17)  

a. El amanecer {es/*está} a las ocho.  
   The dawn {isER /iESTAR} at the eight  
   ‘Dawn is at eight o’clock’

b. El concierto {es/*está} mañana.  
   The concert {isER /iESTAR} tomorrow  
   ‘The concert is tomorrow.’

At first glance, the examples in (17) do not perhaps fit under the traditional description. In the classic literature on *ser* and *estar*, however, the distribution of temporal locatives has been
considered to reflect an immutable characteristic of the individual in question (Porroche, 1990). It has been proposed, that, just like in the case of locatives being inherently temporary, temporal predication is read inherently as a permanent property of an event (Bello, 1955). Under this criterion, *ser* is the expected copula.

When considered in unison, temporal and locative predications seem to escape the temporary vs. permanent distinction. Although it can be argued that the temporal location of an event is a property or inherent characteristic of that event (Leonetti 1994 *i.a.*), it is hard to consider all expressions of physical location as temporary, yet they all can accept *estar*, and in some cases *ser*. In Chapter 2, different alternatives that have been proposed to this challenge will be discussed, while Chapter 3 integrates the data from examples such as (16) and (17) into the main proposal of this study. Nevertheless, locative and temporal prepositional contexts remain one of the more interesting and contentious areas of the *ser* and *estar* phenomenon.

1.3.3 *Ser* and *Estar* with Nominal Complements

In general, only *ser* allows for noun complements. This is perhaps to be expected given that *ser* is the copula of choice for identificational (18) purposes:

(18) a. París  es  la capital  de Francia.
    Paris  is*ser*  the capital  of France
    ‘Paris is the capital of France’

The first president of the States United was George Washington

‘The first president of the United States was George Washington’

The identificational nature of the copula *ser* is made evident in (18). It establishes a one-to-one relationship between the entity *Paris* and the entity *the capital of France*, and *George Washington* and the *first president of the United States*.

There is another context of copula predication in Spanish that has lent support to this view of *ser* being the copula of choice for identification. This context is cleft structures, as the example in (19):

(19)  

a. \{Es/*Está\} Marta quien ha sido ascendida.  
   (It) \{is* /isESTAR\} Marta who has been ascended.  
   ‘It is Marta who has been given a raise’

b. \{Es/*Está\} en Tokio donde será la próxima Olimpiada.  
   (It) \{is* /isESTAR\} in Tokyo where (they)are.FUT the next Olympiad.  
   ‘It is Tokyo where the next Olympic Games will be’

c. \{Es/*Está\} por la mañana cuando nos traen la leche.  
   (It) \{is* /isESTAR\} by the morning when us (they) bring the milk  
   ‘It is in the morning when they deliver our milk’
All cleft structures are strict *ser* contexts in Spanish, and the interpretation of these left-dislocated elements seems to be one where copula choice is ignored, since even when a strictly *estar* adjective is fronted, *ser* remains the only viable option, as can be seen in example (20):

(20) a. {Es/*Está} embarazada lo que Marta está.
    (It) {is*ESTAR} pregnant (it) that Marta *ESTAR*.
    ‘It is pregnant what Marta is’

b. Abierta {es/*está} lo que la puerta está.
    Open {is*ESTAR} (it) that the door *ESTAR*.
    ‘Open is what the door is’

This distribution of copulas is not surprising if the left-dislocated element is considered to be in a position where it needs to be identified. Under this consideration, it would not be surprising that these structures fall under the same copula distribution rules as nominal predicates. It is nevertheless worth mentioning them as the only context in Spanish in which strict *estar* predicates can appear in a *ser* predication.

There are, however, contexts where *estar* can in fact combine with nominal predicates. The examples in (21) are taken from online sources that display this use of *estar*:

(21) a. Rivera estuvo torero en la rueda de prensa.
    Rivera *wasESTAR* bullfighter in the wheel of press
In the examples in (21), the interpretation of the predicate is bound in time and describes a specific, temporal situation. In this manner, these uses of estar follow the traditional temporary interpretation. But, unlike the uses of ser in nominal complementation, estar does not have an identification use here. Instead, the noun is providing access to the qualities traditionally associated with being a bullfighter (21a), or an artist (21b). In these cases of estar with a noun evoking a set of prototypical qualities, the interpretation could change from individual to individual, and from context to context, as the relevant properties could change. For example, in (21a), bullfighters are supposed to be brave, but also cool in the face of danger, and a bit arrogant. In certain contexts, one or the other of these qualities would be the salient one in a predication such as (21a), and the listener will need to be able to pick the right one. One common characteristic of these estar with nominal predications, regardless of interpretation, is that the properties described by the noun are of a temporary nature, tied to a specific event. In this manner, the examples in (21) are semantically very close to the ones observed with adjectives in (11).
There is another case of *estar* with nominal predications in Spanish that has a similar interpretation to the one presented in (21). The examples in (22) illustrate it:

(22) 

a. ¡Marta está cañón!
   
   Marta *es* *cannon*
   
   ‘Marta is very attractive’

b. Estoy boquerón
   
   (I) *am* *anchovy*
   
   ‘I have no money’

The examples in (22) are *estar* contexts with a nominal complement in common use in Spanish. One of the most salient semantic features of these structures is that they are always idiomatic, where the noun used in the predication has an adjectival nature, instead of a purely nominal one (Peñalver-Castillo 2003). As can be seen in (22a), the interpretation is nowhere near a literal one, as a human being—Marta—could never be considered a literal cannon. Likewise, in (22b), I cannot be considered an anchovy in any literal sense. Instead, these predications are given an idiomatic sense. In this case, being a cannon has the metaphorical meaning of being very attractive, and being an anchovy, having little money. As such, these nominal expressions are syntactically Noun Phrases, but semantically they function as attributives, properties of individuals, not as individuals themselves. Given that both looking attractive and being broke can have temporary interpretations, the fact that they combine with *estar* falls under the general
principle of it being the copula of temporary states. Furthermore, the idiomatic meaning of these nouns seems to be attached, at least partially, to their combination with *estar*, since switching the copula precludes the metaphoric interpretation, or at least, results in a less acceptable structure, as seen in the examples in (23):

(23)  a. ??¡Marta es un cañón!
Marta is a cannon
‘Marta is very attractive’

b. Soy un boquerón
I am an anchovy
??‘I have no money’

Furthermore, these nominal predications present yet another problem than the merely idiomatic one. Even when they are applied as a permanent quality of an individual, they cannot occur in a *ser* predication, as (24) illustrates:

(24)  *Mis padres fueron boquerón toda su vida.
My parents were an anchovy all their life.
‘My parents never had any money’
Once again, as (24) shows, the permanent and temporary distinction falls short. Another interesting feature of these *estar* idiomatic structures is that the noun involved in the predication must appear as a bare singular noun. The examples in (25) show how the noun seems to ignore number agreement with the subject and needs to appear without any determiner. Although this syntactic constraint is not relevant at this point, it is worth mentioning and keeping in mind, as it will become relevant in following chapters.

(25)  
\[ a. \quad \text{¡Marta está \{un/el/aquel\} cañón!} \]
\[ \text{Marta \textit{is ESTAR} \{a/the/that\} cannon} \]
\[ \text{‘Marta is very attractive’} \]

b.  
\[ \text{¡Marta y Ana están \{cañón/cannons\}!} \]
\[ \text{Marta and Ana \textit{are ESTAR} \{cannon/cannons\}} \]
\[ \text{‘Marta and Ana are very attractive’} \]

The structures presented in (25) are very close to the ones in (22). In both cases, the nouns involved in the predication provide the meaning more commonly associated with an adjective, \textit{i.e.}, a set of properties that the subject of the predication possesses. The interpretation, however, is not as straightforward in cases like (21) as in the idiomatic (22). One view of the interpretation of idioms (Jackendoff, 1995) holds that the idiomatic reference is one that an individual has in their lexicon. Thus, an interpretation of the word *cañón* where a woman who is a cannon is very attractive, and *boquerón* indicating, besides a type of fish, also a state of being broke would be
stored in a speaker’s mental lexicon. As such, the interpretation of the sentences in (22) would be straightforward, or, at least, as straightforward as a non-idiomatic one. Cases like (21) could be considered as halfway cases, since their meaning can be interpreted from their components—provided one has access to the prototypical attributes associated with the noun of the predication—while a literal reading of the sentence will provide a wrong interpretation.

As we have just seen, there are some interesting and illuminating structures involving estar and noun complements. These particular uses have received less attention than most in the literature, despite the fact that they could offer some critical insights into the ser and estar distinction. Cases of estar with a nominal complement will be more closely examined in Chapter 2.

1.3.4 Ser and Estar with Verbal Complements. Gerunds and Participles.

In the case of verbal complements, both ser and estar can function as auxiliary verbs in Spanish, generating the Passive (26a) and Progressive (26b) constructions respectively. These constructions combine either the ser or estar copula with a Participle or a Gerund form of a verb, respectively. The examples in (26) examine these two types of predicates in more detail:

(26) a. Mario fue detenido por la policía.

Mario was\textsubscript{SER} arrested by the police

‘Mario was arrested by the police’

b. Mario estaba rompiendo los cristales de una tienda.

Mario was\textsubscript{ESTAR} breaking the glasses of a store
‘Mario was breaking a store’s window’

In the case of predications involving verbal Participles, both *ser* and *estar* can be found. The examples in (27) display this distribution:

(27) a. Pepe fue detenido ayer por la noche.

Pepe **was** *ser* arrested yesterday by the night

‘Pepe was arrested last night’

b. El camino está bloqueado

The path **is** *estar* blocked

‘The path is blocked’

The example in (27a), is a purely Passive construction—Pepe is the semantic Patient of an event of getting arrested, where the Agent is left undetermined. As it was previously mentioned, these classic Passive structures always involve the copula *ser*. In the case of (27b), the Participle is now involved not in a Passive predication, but in a Resultative one (Embick, 2004; Marín, 2010). In essence, the Participle in (27b) is functioning much like a de-verbal adjective, simply describing the state that an individual has attained by virtue of an event—in this case, the path is now in a state of being blocked after someone or something blocked it. As Kratzer (2012) suggests, as long as the Participle involved can denote a state, it is possible to construct both a Passive and a Resultative construction. Building on the examples in (27), we can attest to this possibility, as can be seen in the examples in (28):
Leonetti (1994) suggests that this dual distribution of Participles is due to *estar* providing a Resultative or Perfective interpretation to the predicate, that Passive constructions with *ser* lack. It has been proposed that Auxiliaries provide Tense, Mood and Aspectual information to a predication (Hengelved 1986, Porroche 1990, Lema 1992 for an analysis on Spanish). As such, this would fit an analysis where *estar*, by virtue of providing a Perfective Aspect to the predication, also denotes a temporary one.

Turning our attention now to predications involving Gerunds, *estar* is used as an auxiliary to construct the Progressive (29) in Spanish.

(29) a.  Carlos está estudiando en la biblioteca.

      Carlos is*ESTAR* studying in the library

      ‘Carlos is*ESTAR* studying at the library’

b.  Estaba lloviendo.
(It) was ESTAR raining’

‘It was raining’

This is perhaps one of the least problematic areas for copula distribution in Spanish, as there are no known deviations from the norm. Furthermore, if *estar* is considered the copula involved in temporary predications, it falls naturally that Progressive Aspect would employ this copula as its auxiliary of choice. Going back as far as Marchand (1955), Progressive Aspect is defined as having, among other properties, a denotation of a change of state in an individual, thus providing what Roby (2009) describes as a “non-homogeneous temporal reading”. In this sense, a Progressive construction will always indicate a temporal one, where the individual has moved from not having a certain property to actually having it. In (29a), Carlos has moved from a state of not being studying at the library to a different state in which he is. Thus, *estar* is predicted to be the favored copula in these contexts. Once again, it is worth noting that there are no known cases of *ser* combining with Gerunds in Spanish.

1.4 Problems with classical accounts

In the previous sections, all possible syntactic contexts of *ser* and *estar* predication have been examined and explained in terms of *ser* providing a permanent interpretation, and *estar* a temporary one. This classical interpretation is not without problems, however. Although the temporality distinction for *ser* and *estar* is the most widespread account, and it does account for the majority of the data, it is also not difficult to find examples that are apparent contradictions to
this approach. This section focuses on the problematic cases of *ser* and *estar*, following the same syntactic contexts used in the previous section.

1.4.1 Adjectival Predication

In the case of adjectival predication, it is not hard to find counter-examples to the traditional view. The following examples in (30) provide instances of adjectives denoting temporary properties of individuals that nevertheless appear only with *ser*:

(30)  

a. Marta es/*está española.  

Marta \(i^\text{SER} /i^\text{ESTAR}\) Spanish  

‘Marta is Spanish’

b. Marta es/*está musulmana.  

Marta \(i^\text{SER} /i^\text{ESTAR}\) Muslim  

‘Marta is a Muslim’

In (30a), Marta being Spanish does not describe a necessarily permanent situation, as people can acquire a new nationality and lose their original one. This is also not a rare occurrence, so it should not be such an alien concept that would prevent temporary readings of adjectives describing nationalities. Nevertheless, the condition of being Spanish never accepts *estar*.

Similarly, in (30b), the condition of being a Muslim is not necessarily permanent. But, even in cases where we are describing a person who is a Muslim after having converted, *estar* is not an option. One possible explanation to this surprising distribution is that the concept of changing
one’s religion or nationality is relatively new, and perhaps not as simple as changing one’s location. Perhaps this is a context in which social and political advances have not caught on with language yet.

One further consequence of considering *estar* only as denoting temporary properties of individuals is that this requires a previous knowledge of said individual. As such, it should not be possible to have an individual that has never been encountered before as the Subject of an *estar* predication. The example in (31) shows how this prediction turns out to be wrong. In (31), even though we just met the individual Manuel, we can make observations about supposedly temporary conditions even though we do not have access to any previous state that he might have been in.


   Yesterday (I) met to Manuel. (He) *está* enough fat.

   ‘I met Manuel yesterday. He is pretty fat.’

Taking the previous example even further, Querido (1976) not only claims that *estar* can be used to describe situations that are experienced for the first time, but goes as far as claiming that, in fact, *estar* is the default copula that should be used when in a novel situation. His classic example describes a situation where a botanist discovers a new type of tree, deep in the Amazonian jungle. Upon observing that the tree has red leaves, the botanist—Querido claims—would use *estar* to describe the color of the leaves, as shown in example (32), from Querido:


   Yesterday (I) recognized a tree. *Está* red.

   ‘I recognized a tree yesterday. It is red.’
Las hojas de esta nueva especie de árbol {son/están} rojas.

The leaves of this new species of tree {areSER/areESTAR} red

‘The leaves on this new tree species are red.

It is unclear why this should be the case. Naturally, this is counter-intuitive with estar combining with temporary predicates, since, in the absence of further evidence, the botanist has no way of knowing whether the color of the leaves is permanent or not. It is worth noting that the example presented in (32) is one where native speakers of Spanish do not seem to agree.

There is another type of construction with estar and adjectival complements that presents a challenge to the traditional accounts; the evidential uses of estar. In this case, estar is involved in predications (33) that cannot be considered temporal:

(33) a. Este sushi está estupendo.

This sushi is ESTAR stupendous

‘This sushi is fantastic’

b. La habitación del hotel estaba grande.

The room of the hotel was ESTAR big

‘The hotel room was big’
These examples are *estar* contexts where the predication does not imply any kind of temporary condition. There are, however, important differences between (33) and classic counterexamples such as *muerto* ‘dead’. As a first difference, the adjectives involved in this type of predication are not *estar* exclusive. In fact, even in the same contexts where (33) would be acceptable, the *ser* variant could also be produced, as (34) illustrates:

\[(34)\]

\[\begin{align*} 
a. \quad & \text{Este sushi es estupendo.} \\
    & \text{This sushi is}{}^{\text{ser}} \text{ stupendous} \\
    & \text{‘This sushi is fantastic’} \\
\end{align*}\]

\[\begin{align*} 
b. \quad & \text{La habitación del hotel era grande.} \\
    & \text{The room of-the hotel was}{}^{\text{ser}} \text{ big} \\
    & \text{‘The hotel room was big’} \\
\end{align*}\]

One question that immediately jumps to mind is what is the difference between (34a) and (33a). Both the *ser* and *estar* versions of the sentence are very close in meaning. However, it is clear that both predications in (33) denote permanent characteristics of the sushi and the hotel room, respectively. In the case of (33a), the fact that the entity involved in the predication is a determinate piece of sushi favors an interpretation where it has always been fantastic, given how specific meals are inherently either fantastic or not. In the case of (33b), the non-temporary interpretation is not simply the most plausible one, but the only one that seems possible. A hotel room that changes size seems more like a thought exercise in quantum mechanics than anything
that could be possible in the real world. It seems clear that the traditional accounts are not capable of explaining the data presented here.

Additionally, *estar* is not the only source of problems in the distribution of copulas and adjectives for traditional accounts. *Ser* also presents some problems. The following examples show how *ser* can generate temporary predicates.

(35)  

a. Pepe {fue/*estuvo} rico y famoso en su juventud.  
   Pepe {wasSER/wasSER} rich and famous in his youth  
   ‘Pepe was rich and famous when he was young’

b. La falda de María {es/*está} roja.  
   The skirt of Maria {isSER/isSER} red  
   ‘Maria’s skirt is red’

The examples in (35) all involve predicates of a temporary nature – Pepe’s fame and fortune is confined to his youth, and it could be expected that María can dye her skirt a different color quite easily. Yet, these are both cases where an *estar* predication would be considered unacceptable. There is an acceptable interpretation for (35b) with *estar*, where the skirt is not dyed red, but rather dirty with some sort of red substance. Such examples, where a natural process is behind the change in color and result in acceptable *estar* predication will be examined in detail in Chapter 3 and are the focus of the experiment described in Chapter 4. Nevertheless, an
interpretation where the skirt has been dyed red, which is the intended interpretation in (4b), is unacceptable in Spanish with *estar*.

### 1.4.2 Problems in Prepositional Contexts

Prepositional predication, while more stable in the distribution of copulas than adjectival contexts, also provides some data that seems to contravene the temporary vs. permanent account. One such prepositional context that can help in the characterization of the difference between *ser* and *estar*, can be seen in the examples in (36):

(36)  

a. Marta *es* profesora.  

Marta *is* teacher  

‘Marta is a teacher’

b. Marta *está* de profesora.  

Marta *is* of teacher.  

‘Marta is working as a teacher’

The interpretation of these examples, as Camacho (2012) points out, is perhaps subtle and easy to miss at first glance, but very illuminating on the *ser* and *estar* distinction. In (36a), the implication is that Marta is considered a teacher because she has had the training and has obtained all the necessary licenses required to work as a teacher. In semantic terms, it could be said that Marta has acquired the inherent quality of being a teacher. In (36b), all that is being predicated about Marta is that she is currently working as a teacher. It is not necessarily true that
she has her license or that she has received all the formal training usually associated with a
teacher. This presents a contrast with (36a), since in the former we have no information about
whether Marta is currently working as a teacher or not. These two implicatures—Marta might
not be working as a teacher in (36a), and Marta might not be a teacher in (36b)—can be easily
attested in the examples provided in (37):

(37) a. Marta es profesora, pero ahora trabaja de camarera.
   Marta is SER teacher but now works of waitress
   ‘Marta is a teacher, but she is currently working as a server.’

b. Marta está de profesora, pero en realidad es médico.
   Marta is ESTAR of teacher but in reality is doctor
   ‘Marta is working as a teacher, but she is actually a doctor’

Both examples in (37) depict temporary qualities of an individual, but have different
implications. Any theoretical approach to ser and estar needs to be able to account for these
different interpretations.

Locative prepositional contexts are also a source of evidence against estar as a copula of
temporary predicates. The example in (38) shows this:

(38) París está en Francia.
   Paris isESTAR in France
‘Paris is in France’

In (38), it is hard to imagine a situation in which Paris ceases to occupy its present geographical location. The example in (38) describes a situation in which a consensus regarding the temporary or permanent properties of the predicate can be expected. And yet, sentences like the one presented in example (38) are strict estar predicates.

1.4.3 Problems with nominal complements

Just as it happens with adjectival predication, nominal predication does not seem to always follow the traditional accounts provided in the literature. It is possible to find cases of ser nominal predications that do not involve permanent properties, as in the example presented in (39):

(39)   a. Marta es una chica.
       Marta is a girl
       ‘Marta is a girl’

   b. Marta es {una madre/una abogada}.
       Marta is {a mother/a lawyer}
       ‘Marta is {a mother/a lawyer}’

Marta being a girl in (39a) is similar to some previously seen cases, such as nationality or religious affiliation from (33), where our encyclopedic knowledge of the world and linguistic
behavior differ. Nowadays, gender is not as stable a property as it used to be, but perhaps it is too soon for language to reflect this, so the exclusive *ser* nature of this type of predication is not entirely surprising and could be subsumed under the traditional paradigm of inherent properties. Example (39b) presents a different situation. Marta’s being a lawyer or her being a mother cannot strictly be considered permanent or inherent, as people switch careers and children die on occasions. These are properties that have a more transient nature than being a girl, and thus, predicted to be *estar* predications, at least in certain cases. And yet they are not. This distribution is, once again, one that is not predicted by the permanent and temporary distinction. This particularity has led to a generalization frequently used in L2 teaching that considers that only *ser* combines with NPs in Spanish (Zayas-Bazán *et al*, 2015).

1.4.4 Problems with Participles

As was pointed out in section 1.2.2, Participles follow the same rules as adjectives, provided they denote a state and an event boundary (Kratzer, 2012). However, we can find Participles in Spanish, that, despite denoting states of an individual and being bound by an event, do not combine with *estar*, but rather *ser*. The example in (40) shows this behavior:

(40) En los años 80, Vanilla Ice {era/*estaba} muy conocido.

In the years 80, Vanilla Ice {was*ESTAR/estar} very known

‘In the 80s, Vanilla Ice was very well-known.'
This example is one where the predication is clearly temporary and determines a state of an individual. And yet, it is a *ser* exclusive context. Once again, a traditional account based on the transient nature of *estará* cannot properly account for (40). Given the way the predicate is constructed, it clearly has temporal boundaries, and yet, *ser* is the only option available. It seems that in the case of Participles, whether they are involved in a Passive or a Resultative construction, once again the classic theories of temporary vs. permanent cannot account for the observed distribution of the copulas in Spanish.

1.5 Conclusions

This chapter has provided a detailed exposition of the contexts where *ser* and *estará* can be found in Spanish, as well as a general description of the traditional view of how to account for their distribution. In this manner, we now have a basis on which to move on to more modern descriptions, knowing what the classic problematic areas are. This chapter has shown how a theoretical framework that bases its distinction of the copula dilemma in Spanish on inherent, permanent properties for *ser* and temporary, transient ones for *estará* provides an explanation for much of the data, but falls short in too many ways to be considered comprehensive. The array of cases that cannot be accounted for with the temporary vs. permanent distinction sheds light on the fact that, as explained in the immense majority of L2 handbooks, *ser* and *estará* are a source of frustration for learners of Spanish.
This chapter has also provided an overview of all the contexts where both copulas can be found in Spanish. Table 1 summarizes the distribution of the kind of predications that both *ser* and *estar* can occur with.

<table>
<thead>
<tr>
<th>Expression</th>
<th><em>SER</em></th>
<th><em>ESTAR</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>DETERMINER PHRASE</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>NOUN PHRASE</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ADJECTIVAL PHRASE</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>PREPOSITIONAL PHRASE</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>PARTICIPLES</td>
<td>✓ (Passive)</td>
<td>✓ (Stative)</td>
</tr>
<tr>
<td>GERUNDS</td>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 1. Distribution of *ser* and *estar*

In Chapter 2, I expand my analysis to the more modern approaches that have strived to describe copula distribution in Spanish. These newer theoretical approaches will provide a wider knowledge base, which will be used to construct a list of the insights, strengths and weaknesses of each. By observing how different authors handle the challenges of *ser* and *estar*, it is possible to identify the elements that should be in a comprehensive account, which is the objective of Chapter 3.
Chapter 2: Previous Theoretical Accounts

2.1 Introduction

The previous chapter covered the distribution of both copulas, *ser* and *estar*, in all possible contexts in Spanish. This distribution was given a brief explanation along the lines of a classical temporary vs permanent distinction, although the weaknesses of this analysis were also revealed. Despite there being many contexts where a temporary vs. permanent distinction is a valid explanation, there are also many were this is not the case. This second chapter provides an in-depth approach to several of the more popular and influential theoretical accounts that have been put forward to explain the data found in Spanish. These include some that are direct descendants of the classical one, as well as some innovative accounts that have been proposed in recent years. As can be expected, given how central the topic of copula distribution is in Spanish, there have been plenty of approaches to the topic, from all perspectives—syntactic, semantic, pragmatic, etc.—and to this day, it remains a fertile field of study of Hispanic Linguistics.

Finally, to close the chapter, there is a section dealing with the ideas and observations that need to be included in any theory that attempts to explain the *ser* and *estar* dichotomy in Spanish.

2.2 Stage Level and Individual Level

2.2.1 Stage and Individual Level Predicates

One of the theoretical approaches that has attempted to expand on the temporary and permanent distinction of the Spanish copulas has centered on the Individual Level (IL) and Stage Level (SL)
distinction (Carlson, 1977; Kratzer, 1995; Diesing, 1988, 1992; Escandell & Leonetti 2002). The terms Individual Level and Stage Level Predicates date back to Carlson (1977), and were later refined by Kratzer (1995). In general terms, an Individual Level (IL) predicate is one that is stative and determines an essential, time-stable property of an individual. Two typical examples of individual-level predication are given in (41):

(41)  
   a. John knows French.  
   b. Mary is altruistic.

A Stage Level (SL) predicate is episodic in nature, and expresses a property of a stage of an individual. Typical examples are shown in (42)

(42)  
   a. John is speaking French.  
   b. Mary is available.

The concepts of SL and IL predicates have been developed and modified, and their definition has been refined. Carlson (1977) adopts Milsark’s (1974) ideas and his observations on predicate restrictions as a springboard for his own analysis. Milsark had divided adjectives in two groups; one that describes properties that are possessed habitually, and another that describes states. From these observations, Carlson (1977) says that “those [predicates] referred to by Milsark as ‘states’ […] will be predicates that denote sets of stages; the others, like intelligent, will be represented as sets of individuals”. Or, as he formalizes it, ‘states’ are predicated of a particular
stage ‘y’ related (R) to an individual (43a), while ‘property descriptive’ predicates directly introduce characteristics of the individual (43b).

(43)  a.  *John is available:*  \( \exists y [R(y, j) \& A'(y)] \)

b.  *John is intelligent:*  \( \lambda PP \{j\} (I) = I(j) \)

In (43a), there is a stage \( y \), that is related (\( R \)) to the individual John (\( j \)), and that stage \( y \) has the property of availability (\( A' \)). In (43b), given the properties of the individual John (\( PP \{j\} \)) and the property of being intelligent (\( I \)), John has that property (\( I(j) \)). This is the classic definition of Stage-Level and Individual-Level predicates, as presented by Milsark (1974) and reviewed by Carlson (1977). A SL predicate, like (43a), defines the properties of a stage, while an IL predicate defines the properties of an individual. One of the more popular revisions was proposed by Kratzer (1989). In her analysis, she treats SL predicates as including a spatio-temporal or event variable that can be bound by a temporal restrictor; IL predicates, on the other hand, lack this variable. In other words, Kratzer treats SL predicates as Davidsonian predicates, that is, predicates with an event that is part of the argumental structure of the verb. In IL predicates, there is no such variable. Each language would be able to determine idiosyncratically which predicates have this event variable and which would not.

(44)  a.  \( \lambda x \lambda L[P(x, L)] \)  

b.  \( \lambda x[P(x)] \)  

<table>
<thead>
<tr>
<th></th>
<th>Stage Level</th>
<th>Individual Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In example (44a), we can predicate the property $P$ of the individual $x$, but there is also an event argument $L$, which is what allows for temporal restriction of the SL predicate, e.g. *Mary was pregnant last year*, which is not available in IL predicates, which lack this event, e.g. *Mary was intelligent last year*. The SL/IL alternation was formulated as the Stage-Level/Individual-Level Hypothesis, which proposes that this distinction is one that is reflected in a multitude of ways cross-linguistically. Fernald (2000) suggests that this distinction is fundamental to the way that humans think about how attributes of individuals are organized, *vis-à-vis* their boundedness to time, given that it is present in many languages.

One of the consequences of the IL and SL predications are what Kratzer (1989) refers to as lifetime effects. Since IL predicates lack an event attached to their interpretation, conjugating the verb in a Past Tense form will generate a reading where the individual that is the Subject of the predication is no longer alive. This lifetime effect can be observed in (45), where the assumption is that Juan is alive.

(45)  a.  #Juan was intelligent.

b.  Juan was nervous.

In example (45a), the fact that Juan is still alive makes the sentence odd. This interpretation goes back to the Milsark’s (1974) observations about adjectives depicting inherent qualities of individuals, which were the ones that Carlson (1977) later identified as IL predicates. Similarly,
example (45b) is a predicate that involves a stage of the individual, and as such, can terminate without the necessary consequence that Juan also dies.

### 2.2.2 Syntactic Implications of the SL/IL hypothesis

Although the SL and IL distinction is mostly concerned with the semantic interpretation of predicates, it has also been extended to have a reflection in the syntactic structure of the predicates in question. Diesing (1988, 1990) takes this distinction and formulates the VP/IP Split Hypothesis. Her hypothesis presupposes that Verb Phrases (VPs) contain a base generated Subject in all cases, even if the Verb is not unaccusative. These Subjects are generated at the Spec VP position, and, in the case of SL predicates (46a), may move to the Spec IP position, leaving a trace behind. In the case of IL predicates (46b), the Subject is base-generated as a PRO, and only in the Spec IP position is then manifested as an overt Subject:

(46)  

\[
\begin{align*}
\text{(46) a.} & \\
\text{b.} & 
\end{align*}
\]
The VP/IP Split Hypothesis has its reflection in the semantic interpretation, according to Kratzer (1995) and Diesing (1990; 1992). Their proposal is that a grammatical element’s ability or inability to appear in the nuclear scope or in the restrictive clause of its logical semantic representation is directly derived from its position in the syntax. This syntax-semantics interface interaction reflects what Schmitt (1992) calls the Mapping Hypothesis, which states that material inside the VP is mapped onto the nuclear scope and material outside of the VP onto the restrictive clause. Nuclear scope is the material that forms the predication that the quantifier ranges over, while the restrictive clause restricts the range of the quantifier. For example, in a sentence like *Most students left*, students would be the restrictive clause of the quantifier *most*, and *left* its nuclear scope. This difference means that overt subjects of SL predicates, which remain in Spec VP or else raise to spec IP, are logically represented as being either within the nuclear scope, if they do not raise, or in the restrictive clause, if they do raise. Overt subjects of IL predicates are always logically represented as being within the restrictive clause, since they originate at Spec IP. The graphs in (46) show this schema. In (46a), the subject of an SL clause always has can be both the restrictive clause if they rise to IP, or nuclear scope if they do not. In (46b), the subject of an IL clause must raise and can only appear as the restrictive clause.

### 2.2.3 The SL/IL distinction and *ser* and *estar*

The distribution of *ser* and *estar*, which had been traditionally interpreted in terms of permanent and temporary properties (see Chapter 1), seems to correlate with IL and SL predicates (Kratzer, 1995). This is not surprising, since the way SL and IL predicates are formulated tend to give the
former a permanent interpretation, and the latter a temporary one. The examples in (47) show how this interpretation applies in Spanish:

(47)  a.  Juan es nervioso.  
      \begin{align*} 
      & \text{Juan is} \text{SER nervous} \\
      & \text{‘Juan is a nervous person.’} 
      \end{align*} 

b.  Juan está nervioso.  
      \begin{align*} 
      & \text{Juan is} \text{ESTAR nervous} \\
      & \text{‘Juan feels nervous.’} 
      \end{align*} 

Example (47a) involves a predication over the individual Juan, indicating that he has the property of being nervous. In (47b), on the other hand, the predication ranges over a stage in the life of the individual Juan, a stage in which Juan had the property of being nervioso ‘nervous’. In the terms of Kratzer’s account (1989), these could be represented as:

(48)  a.  [nervioso’(j)] 

b.  [nervioso’ (j, e)] 

One test that is used to determine the status of a predicate concerning their IL or SL status consists in adding a temporal restrictor to the predication. A SL predicate will accept it, since this restrictor can quantify over the stage $e$, whereas an IL predicate will not, since it lacks such a stage. This test can be applied in a straightforward way: adding a restrictor such as \textit{En Navidad} ‘At Christmas’ to (48b) yields a valid sentence \textit{Juan está nervioso en Navidad} ‘Juan is \text{ESTAR nervous}
nervous at Christmas’, but will produce an odd sentence when combined with (45a) #Juan es nervioso en Navidad ‘Juan is SER nervous at Christmas’

One influential line of thought on the difference between ser and estar, as seen in Diesing (1992) and Lema (1995), for example, assumes that their contrast is a manifestation of the SL and IL predicates distinction. As noted, Diesing (1992) suggests that the SL/IL distinction can be applied to Spanish, in that the distribution of ser and estar correlates with IL and SL predicates respectively. Her analysis is based on the VP/IP Split Hypothesis: She claims the Subject of ser is generated in IP. Since it cannot reconstruct to VP, it can only receive a generic reading. The subject of estar is generated inside the VP, and can then move to IP to receive a generic interpretation. This yields the pattern in (49):

(49)  

a. Un bombero es valiente.  
A fireman is SER brave
‘A fireman is brave’

b. Un bombero está disponible.  
A fireman is ESTAR available
‘A fireman is available’

In example (49a), the only available reading is the generic one—all firemen are brave. In example (49b) however, the interpretation can be both a generic one—all firemen are available—and an existential one—there is one specific fireman, and he happens to be available.
The SL/IL distinction is, to this day (Lema, 1995; Zagona, 2009; Marín, 2009, *inter alia*), one widely used account for copula distribution in Spanish. Its popularity can be seen in how the distribution of *ser* and *estar* is typically explained to learners of Spanish—*ser* is for permanent properties, *estar* for temporary ones. There are, however, many authors that have challenged this view. Marín (2009), for example, states “contrary to what is traditionally assumed, the compatibility with *ser* or *estar* is not a defining diagnosis for determining the IL or SL nature of Spanish adjectives”. Some of the counter-examples that he cites include classic problematic adjectives such as those in (50):

(50)  a. Napeoleón está muerto.

Napoleon is *ESTAR* dead.

‘Napoleon is dead’

b. En otoño, las hojas de los árboles son rojas.

In fall, the leaves of the trees are *SER* red.

‘In the fall, tree leaves are red.’

In the case of (50a), the predicate *muerto* ‘dead’ is a permanent, IL one. It is not logical to think of Napoleon’s condition of being dead as temporary, and yet the adjective *muerto* must combine with *estar*. Likewise, tree leaves being red during the fall in (50b) is a transitory condition, a Stage Level predicate of said leaves. If the temporal restrictor test previously mentioned is applied in this case, it can be seen that these predictions are confirmed. IL will not accept such a
restrictor, whereas SL will. In our previous example, the temporal restrictor that (50b) has, *en otoño* ‘in the fall’, does not affect the acceptability of the sentence. If that restrictor were to be applied to (50a), the sentence becomes extremely odd #*En otoño, Napoleón está muerto* ‘In the fall, Napoleon is dead’. This last sentence is only acceptable if the expression *in the fall* does not denote a temporal interval, but rather a specific date or point in time. That is not the case in (50b), which can have a reading where tree leaves are red in the fall, and then they change their color. The temporal restrictor test can be applied in all cases to determine whether a predicate denotes a Stage or Individual Level.

Nevertheless, the IL predicate in (50a) combines with *estar*—and must do so, in fact—and the SL one in (50b) is perfectly acceptable with *ser*. These counter-examples to the SL/IL hypothesis seem to indicate that, although it can account for a large portion of the distribution of copular predicates in Spanish, it does not suffice to explain the whole of the phenomenon.

Yet another area where the SL/IL distinction fails as an indicator of copula choice in Spanish comes from some *estar* predicates that Spanish does not allow, and yet are predicted by this framework. Two such examples are provided in (51):

(51) a. María era rica/famosa.

   *María was*<sup>SER</sup> famous/rich

   ‘María was famous/rich’

b. María era inteligente/divertida.
María was \textit{ser} intelligent/fun

‘María was intelligent/fun’

Logic dictates that ‘famous’ and ‘rich’ are predicates that can describe both an essential quality of an individual - IL predicate – or a stage of an individual – SL predicate. Moreover, the usual tests that are employed point towards this interpretation. For example, Kratzer points out the lifetime-effect of IL predicates. A predicate over an individual must apply to its complete lifetime, since there is no event variable to temporally restrict said predicate. The examples in (51) can be used to test this lifetime effect, both of them being in Past tense. Since the Past tense used presupposes that the predication is no longer operative in the present, IL predicates, lacking an event variable, will necessarily need to lead to the conclusion that the subject of the predication is dead.

This is however not the case. Example (51a) is underspecified regarding the current situation of the subject \textit{vis-à-vis} their being alive or dead. In other words, (51a) does not give us any information about whether María is still alive. Example (51b), on the other hand, is a strange statement if María is still alive. A such, it must be concluded that using the distribution of \textit{ser} and \textit{estar} as a diagnostic for IL and SL predicates respectively is not applicable, since there are predicates that pass the traditional SL tests that do not appear in \textit{estar} structures.

Furthermore, Diesing (1992) also notes that there are contrasts in the interpretation of predicates that occur with some kinds of indefinite subjects. Singular indefinites as subjects of a SL
predicate accept both a specific, as well as a generic reading. An IL predicate, on the other hand, will only accept a generic interpretation. The examples in (50) include an indefinite subject, and a *rico* ‘rich’ and *famoso* ‘famous’ predicate each:

(52)  
a. Un príncipe saudí es rico.  
‘A Saudi prince *is* rich’

b. Un ganador del Oscar es famoso.  
‘An Oscar winner *is* famous’

Both examples in (52) can have a generic reading – all Saudi princes are rich and all Oscar winners are famous—as well as an existential one – there is one particular Saudi prince who is rich, and one particular Oscar winner who is famous. These dual interpretations lead us to consider these predicates as stages of an individual. Following the hypothesis that identifies SL predicates with *estar*, we should be able to, at least, find evidence of *rico* and *famoso* being adjectives that can combine with both *ser* and *estar*. This prediction turns out to be wrong.

### 2.2.4 Stage-Level, Individual-Level and Events

Maienborn (2005) also presents an interesting set of examples in trying to show that the *ser* and *estar* distinction is not based on a SL/IL criterion. If *estar* predicates include a Davidsonian eventuality, they should be able to function with locative predicates. Yet, contrary to this prediction, the examples Maienborn provides in (53) are unacceptable:
If *estar* introduced an eventuality argument—being on the table in (53a)—, a locative expression should be able to modify it, expressing that in example (53a) there is, as Maienborn puts it, a state of the shirt being wet, and that this state is physically located on the chair in question. In other words, we should be able to interpret (53a) as *when you put the shirt on the chair, it becomes wet*. Maienborn claims that the unacceptability this interpretation of (53), which patterns with the one observed for the *ser* predicates, implies that they do not include a Davidsonian eventuality, and thus, cannot indicate a semantic distinction based on SL and IL predication.

In the case of evidential uses of *estar* predicates, the SL/IL distinction is unequipped to account for the data. This is only to be expected, since this framework is based on the eventive structure of the predicates, which accounts for only time-based distinctions. As has been previously pointed out, many of the evidential uses of *estar* escape these type of temporal considerations altogether, and depict situations where there is no logical way to think that they might ever change. Consider for instance the examples in (54):
These evidential uses present an undeniable challenge not only to proposals based on the SL/IL distinction, but to all theoretical accounts. In this particular example, the subject of the predication—the room—is one that cannot logically have an event attached to its being big, since rooms don’t change in size. If we apply the aforementioned temporal restrictor test and add an expression such as hoy ‘today’, as in (54b), the interpretation of the sentence becomes odd. There is an alternative interpretation of the sentences (54) that does not result in (54b) being unacceptable—if the room has undergone some renovation and it has been enlarged, both sentences in (54) can be applied, but in this case, this use of estar would fall in the temporal category—it would be a similar use of the one found in cases such as La habitación está limpia ‘The room is ESTAR clean’, which are SL predicates. But, the fact that (54a) is acceptable in cases where the room’s size has never been altered, makes these evidential uses escape the SL/IL interpretation. It could be argued that perhaps evidential uses of estar are cases of certain adjectives being marked in the lexicon as being able to combine with estar, regardless of any
other consideration. This particular interpretation assumes that, in these cases, both *ser* and *estar* versions of the same sentence would have the same interpretation. This is however not the case. Evidential uses of *estar*, by virtue of the greater degree of involvement on the part of the speaker, carry with them certain implicatures that cannot be cancelled:

(55)  

a.  
Este vino es excelente, pero a mí no me gusta.  
this wine is*ser* excellent, but to to me not me like  
‘This wine is excellent, but I don’t like it’

b.  
Este vino está excelente, pero a mí no me gusta.  
this wine is*estar* excellent, but to to me not me like  
‘This wine is excellent, but I don’t like it’

The difference observed between (55a) and (55b) point to interpreting evidential uses of *estar* and their *ser* counterparts as having different interpretations, and not simply as cases of lexical representation.

Despite their problems, the SL/IL approach to copula distribution in Spanish remains one of the more popular ones, and many of the later proposals have attempted to build on its momentum. In the field of language pedagogy, for example, the standard remains to this day to teach L2 learners of Spanish about the difference between *ser* and *estar* still in terms that closely resemble the concepts presented in the SL/IL school of thought.
2.3. Aspectual Theories

There have been several authors (Fernández Leborans, 1995; Roby, 2009; Camacho, 2012) who have tried to extend the temporal consequences of the SL/IL into an aspectual framework that attempts to explain the distribution of *ser* and *estar*. As an alternative to the SL/IL conception of the *ser/estar* dichotomy, these authors trace the root of the difference in the distribution of the copulas to aspect. Specifically, *estar* is assumed to be a marked [+Perfective] copula, and *ser* either [-Perfective], (Fernández Leborans, 1995; Camacho, 2012), or as a default, unmarked copula (Roby, 2009). The aspectual analyses tend to focus more on the nature of the event, and predict interactions with other aspectual markings (such as the preterite/imperfect tense marking).

For Fernández Leborans (1995), *estar* requires that its predicate possess an aspectual feature. In his analysis, *estar* denotes an event, following Pustejovsky’s (1990) typology of aspect, that is composed of a transition (T) and an end state (ES). In this sense, the sentences in (56) can be taken as an example:

(56)  

a. Juan está feliz  
    Juan *i*ESTAR happy  
    ‘Juan is happy’

b. Juan es feliz  
    Juan *i*SER happy  
    ‘Juan is happy’
The predication in (56a) includes both a couple of states—one of happiness and one of not-happiness—and a transition between them, that the subject Juan has gone through. The sentence in (56b), on the other hand, only encodes one single state, that the subject Juan is in—happiness. The end state ES, however, should not always be conceived as the direct consequence of a resultative event, and, for example, estar soltero ‘being ESTAR single’ isn’t necessarily the result of ‘becoming single’.

Roby (2009) and Camacho (2012) also propose theoretical accounts for the ser and estar based on the aspectual properties of the predicates associated with either copula. Both authors take the SL/IL distinction as their starting point, and try to develop that concept further to refine the intuitions it purports and resolve the issues that plagued it. Camacho assumes that “estar agrees with its complements in an inchoative feature, which encodes the beginning point of an event”. His proposal is based on estar selecting for the beginning of a boundary of a state. For example, going back to our previous example in (56), estar is selecting for the inception of the state of Juan being happy. Camacho proposes then that estar includes an uninterpretable inchoative feature [INCH] that must be satisfied by the predicate in order to validate the structure. Regarding ser, Camacho, following Zagona (2010), considers it to be unmarked for aspect. Essentially, under Camacho’s considerations, the key feature in order to determine the distribution of both copulas is how they interact with event boundaries: while Zagona proposes that estar requires a transition from a previous event to the one determined by the estar predication—in (54), estar indicates that the individual Juan has moved from a previous, non-happy state, to a happy one—
Camacho formulates that *estar* does not require the identification of a previous state, and rather it simply selects for the event boundary that indicates the beginning of the state, for example, being happy in (54). He formalizes this feature as [INCH], or Inchoative. And, just as in Zagona’s proposal, *ser* lacks any requirement of the sort, and it simply focuses on the state determined by the event itself. Roby (2009) proposes a similar account for *ser* and *estar*, although in his case, the aspectual feature that best captures the differences is perfectivity. *Estar* is a [+ Perfective], whereas *ser* is [− Perfective]. This is a similar proposal to Camacho (2012), and it is based on the relative importance of event boundaries in the predicates that each of the copula select.

### 2.4 Syntactic Accounts

#### 2.4.1 Zagona’s Syntactic/Aspectual Account

Zagona (2009, 2010) has proposed one of the more influential theoretical attempts to capture the *ser* and *estar* distinction based on the syntactic features of both copulas. According to Zagona the *ser/estar* distinction is a syntactic process, under the minimalist framework, that generates the spell-out of a functional verb BE as either *ser* or *estar*. Under her interpretation, there are no semantic differences between the two verbs, and their selectional properties are reduced to a purely structural difference. In her own words, *ser* and *estar* are “a byproduct of a difference in their formal features. Briefly stated, I argue that *estar* has an uninterpretable prepositional feature [...] checked by a preposition in the complement of *estar.*” In this manner, her analysis is in line with not only many interpretations based purely on intuition, but also more systematic approaches (*inter alia*, Luján, 1981; Clements, 1988; Fernández Leborans, 1999; Escandell-
Vidal & Leonetti, 2002; Marín, 2004, 2010; Camacho, 2012), which consider that estar provides an additional layer of meaning that ser does not.

Zagona’s proposal states that estar, which is generated at the v Phrase, has an uninterpretable prepositional feature, [uP], which must be checked by a complement. In this sense, the vP will serve as a verbalizer (Gallego & Uriagereka, 2009), transforming the root of the copula into a verb, as shown in (57):

\[(57) \quad \text{estar: } [v[uP] \ldots]\]

The complement matching the [uP] feature has two properties:

1. It must also be prepositional in nature.
2. It cannot contain a certain lexical-aspectual content, which Zagona characterizes as a Path.

In her proposal, estar always denotes the presence of a Locative element, whereas ser is the elsewhere copula. Extending this analysis to all uses of ser and estar, her proposal can be distilled into the notion that the distribution of estar is explained in terms of the syntactic structure reflecting certain aspectual properties, whereas the syntactic structure of ser is a consequence of the absence of those aspectual features. One context in which it is easy to see
how Zagona’s analysis fits the data is in the distribution of *ser and *estar with DPs. The following examples in (58), taken from Camacho (2012), show this distribution:

(58)  

a. *Obama está[uP] [DP [N president]]

   Obama iestablish president.

   “Obama is president”

b. Obama está[uP] [P de [DP [N president]]]

   Obama isestablish of president.

   “Obama is president”

As can be seen, in example (58a), the [uP] feature remains unchecked, since the DP cannot satisfy its requirement—DPs lack any prepositional element, overt or covert, that would satisfy the prepositional requirement that [uP] requires. In example (58b), on the other hand, the presence of the preposition *de checks the feature on *estar, and the sentence is no longer acceptable. In this case, Camacho (2012) suggests, the preposition *de is acting as a dummy preposition, and is simply there to satisfy the [uP] feature.

Zagona chooses Locative complements as the initial point of her analysis, and, in this case, she suggests a contrast between prepositions that indicate strict Location (59a) and those that indicate the previously mentioned Path (59b):
(59)  a. Los turistas {están/#son} en Egipto. LOCATION

“The tourists {are/are} in Egypt.”

b. Este regalo {es/#está} para José. PATH

“This gift {is/is} for Jose.”

Zagona’s proposes that Prepositional Phrases that include a Path are headed by an aspectual head, P_{ASP}, which is incompatible with the [uP] feature present in estar. The [uP] is therefore left unchecked. In the case of Location PPs, a non-Path PP checks the [uP] feature of estar.

Consider the examples and their syntactic structure in (60):

(60)  a. Pablo está en casa.

Pablo isESTAR in home

‘Pablo is at home’

b. *Pablo está a casa.

Pablo isESTAR at home

‘Pablo is at home’
In example (60a), the \( uP \) feature in \( está \) gets checked out by the preposition \( en \), whereas in (60b), the aspectual head \( P_{ASP} \) blocks the checking of the feature and the derivation fails.

Zagona’s analysis can also explain the contrast observed between subjects expressing eventive and non-eventive situations. The examples in (61) illustrate this:

(61)  

\begin{itemize}
  \item a. La pelota \( está \) en la mesa.  
      The ball \( iš\text{estar} \) on the table  
      “The ball is on the table”
  \item b. La fiesta \{es/*\( está\} en la discoteca.  
      The party \{iš\text{ser}/*iš\text{estar}\} in the disco  
      “The party is at the disco”
\end{itemize}

Zagona assumes that a non-eventive subject such as the one in example (61a) does not denote a Path, whereas eventive ones, like (61b), do. This results in a distribution where \( estar \) can combine with the former, but not the latter. This interpretation also solves the problems that a SL/IL account has when handling locatives. For example, while the example in (61a) falls under the explanation of \( estar \) defining the properties of a stage in time, locative expressions that denote properties of individuals such as The Sun being at the center of the Solar System, also require the use of \( estar \).
Under Zagona’s view, the temporal properties of each construction — *estar* plus adjective phrases, prepositional phrases or gerunds — are determined by the complement, not by *estar*. When *estar* behaves as an aspectual auxiliary, it takes as complement a Locative Phrase that introduces a Reference time, which is in turn ordered relative to event time: the ordering relation specifies in each case a different aspectual value for the construction (perfect, progressive or prospective). When *estar* behaves as a copula and combines with adjectives, the temporal effects that are associated with *estar* follow from two factors, according to Zagona: the IL/SL distinction, which is encoded in the adjective phrase, and perfective aspect, which links the stages of the adjectival predicate to the Reference time of the clause. This implies that adjectives in copular environments can appear in two different structures: SL adjectives are embedded within an Aspect Phrase and a Locative construed as a temporal location (Reference time), whereas IL adjectives must appear in a context other than the Locative-Aspect structure (notice that this entails that adjectives are always structurally ambiguous in languages that do not display two copulas). In Spanish, the copula is spelled out as *estar* as a consequence of Locative agreement. In the absence of a Locative context, the copula is always spelled out as *ser*. This syntactic account is extended to cover the combinations of copulas with prepositional phrases: Locative agreement is possible when *estar* is followed by locational PPs, whereas it is blocked with directional PPs, since their more complex internal configuration — in particular, the Path component — intervenes between the copula and the Locative phrase; the same holds for the incompatibility of *estar* and eventive nominals as subjects, due to the presence of a Path component in eventive nouns. The central idea is that “the apparent aspectual difference between
the two copulas is not due to a semantic feature of the copulas themselves, but follows indirectly from the properties of the categories they merge with”.

While Zagona’s analysis handles PPs complements in a very straightforward manner, it is, however, not without problems outside of these contexts. One area in which her framework falls short is in cases of adjectival predication that can combine with both *ser* and *estar*. For instance, both *estar nervioso* ‘to be nervous’ and *ser nervioso* ‘to be a nervous person’ are available in Spanish. It is unclear how being a nervous person contains a Path, as she defines it, but being nervous right now does not. Zagona’s analysis does not seem capable of explaining the different interpretations of such cases. Furthermore, the evidential uses of *estar* seen in Chapter 1, such as *El vino está estupendo* ‘The wine is great’ present yet another context where both copulas are possible, also escape her analysis.

### 2.4.2 Gallego & Uriagereka

Another influential proposal, and one that shares many similarities with Zagona’s, is proposed by Gallego & Uriagereka (2011, 2016). In their theoretical account of the distribution of copulas in Spanish, Gallego & Uriagereka argue that “*estar* is *ser* plus an additional element: X, which we have speculated may be prepositional in nature”. This is depicted in (62), for *Alberto is inteligente* ‘Alberto is intelligent’:

\[
(62) \quad [\text{estar}_P \text{ ser } [X_P [X [SC \text{ WP } Y_P ]]])
\]

[where WP = subject, YP = predicate]
[estarP ser [XP X [SC Alberto inteligente] ]]

Just like Zagona and other previously mentioned accounts, Gallego and Uriagereka base the semantics of their proposal on the presupposition that estar has a SL/thetic interpretation. In this context, thetic refers to Kuroda’s (1972) proposal, where he states that a thetic interpretation gives emphasis to the event involving the object, whereas a categorical interpretation assigns a certain property to the object. This thetic/categorical distinction echoes the SL/IL one, in the sense that thetic/SL interpretations of predicates are based around an event argument of the predicate in question, whereas a categorical/IL one is based on temporally pervasive ones. Of course, this gives raise to the same issues that can be found in accounts such as the ones previously mentioned following a SL/IL distinction 2.2.3 of this chapter. While it is true that many estar predicates are indeed SL, this is not a necessity.

Furthermore, Gallego & Uriagereka state in their basic assumption that, from a syntactic viewpoint, estar is more complex than ser. As has been previously shown, estar sentences contain an additional preposition projection, along the lines of the Zagona analysis. The authors’ semantic proposal is that, given the syntactic structure of estar, the predicate can move to a position where it scopes over the subject, grounding it in a particular context, and thus producing the thetic/SL interpretation necessary for a estar context. In other words, by moving to a higher position in the sentence, the predicate essentially scopes over the subject and its semantic properties. For example, in a sentence such as Alberto está nervioso ‘Alberto is nervous’, the
predicate *está nervioso* would scope over *Alberto*, and thus would be a predication of a stage of being nervous, where Alberto is a participant.

These syntactic accounts are, in general, pointing towards a common idea—the copula in Spanish always has the same interpretation, and it is a feature, or features, in its predicates that give rise to either *ser* or *estar*. If the distinction is syntactic in nature, this implies that the two copulas are endowed with different formal features (Zagona, 2012; Camacho, 2012), or that they are inserted in two different syntactic structures (Raposo & Uriagereka, 1995; Camacho, 2015). One reason to explore such syntactic possibilities is the difficulty in finding a valid generalization based on a unitary semantic factor, as was evident in the SL/IL accounts.

One of the areas where syntactic approaches seem to not be able to account for the data that can be found in Spanish is the evidential uses of *estar*. Whether we assume Zagona’s SL nature of the [uP] feature or the thetic/SL level one presented by Gallego & Uriagereka, accounting for examples such as the one in (63), taken from Romeu (2015), presents a challenge:

(63)  El jamón serrano está fenomenal.

The ham serrano *is* *estar* phenomenal.

‘Serrano ham is phenomenal’

While it is true that an evidential reading of (63) involves a thetic judgment (Kuroda, 1972) since it might involve a perceptual report that affirms the existence of an eventuality of a ham being
phenomenal, it is hard to reconcile with a SL reading. Supposedly, and although evaluating the quality of the ham is a subjective operation, the quality of this particular piece of ham is not going to change so that it becomes a product of lesser quality. In this sense, once again an interpretation that is based on temporal qualities, even partially, cannot capture these evidential uses of estar in Spanish.

In general, syntactic accounts like those proposed by Zagona and Gallego&Uriagereka seem to also fall short for some of the same reasons that proposals based on the SL/IL distinction—if the assumption is that the ser and estar distinction in Spanish is based on some sort of temporal or aspectual feature of their complements, those uses of estar that denote a permanent property are always going to fall outside of the analysis.

2.4.3 Romeu (2015)

A more recent proposal is the one presented by Romeu (2015). His proposal is grounded on the nanosyntactic architecture (see Starke, 2005, 2006 for an overview). Romeu assumes a model based on a minimal cartography where every syntactic node in the structure supplies part of the compositional meaning of the whole structure. Romeu’s innovative perspective does not come from the syntactic implementation, but rather from the semantic notions that underlie it. Instead of following the classic distinctions that many of the syntactic proposals employ, Romeu instead assumes that estar predicates denote a state of the individual that is linked to a different state, but, critically, one that the subject of the predication must not necessarily have been in. Ser predicates on the other hand only require that the subject be in a state, with no further
requirements. For example, a sentence like *Alberto es inteligente* ‘Alberto is SER intelligent’ refers only to the current state of intelligence that Alberto has. A sentence like *Alberto está preparado* ‘Alberto is ESTAR ready’ refers both to the present state of Albert as being ready, and another potential state, where Alberto is not ready.

Romeu’s proposal has the main benefit of not imposing any temporal conditions on the copulas, only *estar* needs to be conceptually connected to another, different condition. The following section delves deeper into this particular theoretical framework.

Romeu’s proposal centers on the notion that the distinction between *ser* and *estar* is based on the presence on the syntactic structure of a specific modifier—Conjoint in the case of *ser*, and Disjoint in the case of *estar*. The Conjoint modifier provides the interpretation that an individual is not linked to any other, separate state. For example, in a sentence such as *Juan es inteligente* ‘Juan is SER intelligent’, the *ser* copula requires that the predication only refer to the present state of Juan, without any further implications. The Disjoint modifier, on the other hand, considers the state in which the individual is at the time of the event as being linked to a separate, different state. For example, in a sentence such as *El vaso está vacío* ‘The glass is ESTAR empty’, the use of *estar* indicates that we are connecting the present state of the glass—empty—to another, possible state—full. This interaction between the copulas and their predication and subjects is formalized as shown in (64):
The main claim is that *ser* lexicalizes a stative projection (*Stat* in 64a) together with a Conjoint, whereas *estar* lexicalizes *Stat* with a *Disjoint*. Romeu follows in the footsteps of other authors (Gumiel-Molina *et al.*, 2015, for example) in assuming that the choice of copula depends not only on the properties of the predicate it combines with, but that the copulas themselves have certain properties that condition their combination with certain predicates and subjects.

Perhaps the most important contribution that Romeu brings to the table is to consider the predications involved in both copulas in Spanish as not referring to different states that the subject has had, but rather consider those that could have had. This crucial distinction can be considered an evolution of both the SL/IL and aspectual proposals, in the sense that Romeu’s can account for the lack of that temporal information being present in copular predications in Spanish. Consider the sentence in example (65):
(65) Las casas rojas están vacías.

The houses red are ESTAR empty.

‘The red houses are empty’

The sentence in (65) does not provide any information about the previous state of the red houses, and neither does it presuppose any change in the future. Following a theory where estar requires a strict SL or Perfective aspect interpretation—as proposed by Kratzer (1989) or Camacho (2012), for example—a sentence such as (65) should require the red houses to have been previously occupied, or that they will be in the future. Romeu’s proposal adds another layer of sophistication to previous accounts, and under his view, a estar predication like the one in (65), will be true whether it indicates a true SL predication, or simply one where the houses being not empty is a potential state. This is the step forward that Romeu proposes.

There are certain aspects of the ser and estar distribution, however, that seem to escape Romeu’s proposal. In the case of adjectives such as rico ‘rich’, famoso ‘famoso’ or importante ‘important’, which only combine with ser, Romeu’s account should predict that they would be able to combine with estar. As has been previously mentioned, none of these adjectives actually do:

(66) a. Juan es {rico/famoso/importante}.

Juan is SER {rich famous important}
‘Juan is {rich/famous/important}.’

b. ??Juan está {rico/famoso/importante}.

Juan isESTAR {rich famous important}

‘Juan is {rich/famous/important}.’

In (66a), given the presence of a Conjoint modifier, it is not the case that any of the adjectives provided in the example are connected to a previous state. In this case, they could be said to represent permanent, stable properties of Juan. In the case of (66b), Romeu’s theory does not provide an accurate explanation as to the fact that those sentences are very odd in Spanish. Conceptually, it is not hard to conceive a situation in which expressing that Juan is rich, famous or important is connected to another situation in which he is not. This gap is especially relevant when compared with examples such as the one in (67):

(67)  a. Juan está {nervioso/feliz/deprimido}.

Juan isESTAR {nervous/happy/depressed}

‘Juan is {nervous/happy/depressed}.’

b. *Juan está feliz y (es) famoso.

Juan isESTAR happy and (isSER) famous.

‘Juan is happy and famous.

The fact that the sentences in (67a) are acceptable and commonplace in Spanish, whereas the ones in (66b) are not, seems to indicate the presence of some element that is blocking the estar
interpretation in the latter cases. Further evidence can be seen in coordinated structures such as (67b), where the copula *ser* in the first predicate *feliz* is not enough to validate a coordination with *famoso*, and *estar* is necessary. In the case of Romeu (2105), there is no indication what this blocking element could be. As such, it seems that there are still gaps in his interpretation.

Another point where Romeu’s proposal does not seem to capture the differences in the interpretation between the two copulas is in the evidential cases of *estar*. He examines the same example we saw previously in (68), from Camacho (2012):

(68)  

Este jamón serrano está fenomenal.  

this ham serrano is\textit{estar} phenomenal.  

‘This serrano ham is phenomenal’

In this case, he correctly suggests that the Disjoint state that *estar* requires cannot be a previous one, as it is not logical to think that a particular ham has evolved from a non-phenomenal state to a phenomenal one. Instead, the Disjoint state must be “a non-current one”. Romeu offers the idea that this state that is linked to the current one is a prototypical one for the subject. This proposal, however, creates a situation that can be considered parallel with certain structures with *ser*. For instance, in the following sentence:

(69)  

Este jamón serrano es fenomenal.  

this ham serrano is\textit{ser} phenomenal.
'This serrano ham is phenomenal'

In example (69), the sentence can be interpreted as one in which this particular ham is at least as good as a prototypical ham. As such, it could be argued that both (68) and (69) are in fact cases where the copula is combined with a Disjoint state—which goes against Romeu’s proposal.

2.5. Scalar Accounts

One common thread that many theoretical approaches to the ser and estar share is that they gravitate around a temporal distinction. Whether the key factor is temporary vs. permanent properties, Stage Level or Individual Level predicates, or aspectual considerations, most authors consider that both copulas in Spanish are somehow connected to the temporality of the predicates they combine with. Gumiel-Molina et al (2015) propose an alternative view, based not on temporal properties but on the scalar properties of these predicates. Their analysis only considers adjectival contexts, and propose that the relative/absolute character of adjectives (Kennedy, 2007) is the relevant property underlying their distribution with the copular verbs ser and estar in Spanish.

Gumiel-Molina et al. (2015) center their analysis in the scalar properties of the predicates—adjectives, in this case—that combine with each of the copulas. In this manner, the authors follow in the footsteps of Zagona (2012), and Gallego & Uriagereka (2009, 2016) in considering that ser and estar do not contribute to the semantic interpretation of the sentence, and they are mere spell-out reflexes, simply requiring that their predicates possess certain properties. These
selectional properties that the predicates must possess are articulated, according to Gumiel-Molina et al. around the type of comparison class of the adjectives. Comparison classes are “the implicitly or textually determined sets that relativize the standard for some gradable adjectives” (Bale, 2011). For example, given a sentence with a gradable adjective, such as the one in (70):

(70)  Alex is tall.

The comparison class in this context would be comprised of all the relevant individuals, and Alex’s height is favorably compared to the standard determined by those individuals. In example (68), the comparison class is implicit—in absence of any other information, we could assume that it is made up of all men—but it does not need to be. For example, in cases such as (69), the comparison class is overtly determined to be made up of football players, and Alex is tall when compared to them:

(71)  Alex is tall for a football player.

Gumiel-Molina et al. (2015) establish a connection between the type of comparison class to the type of predicate that ser and estar require in Spanish. They express this connection as “gradable adjectives evaluated with respect to a within-individual comparison class including stages, that is, absolute adjectives, will co-occur with estar. Relative adjectives evaluated with respect to a between-individuals comparison class will surface with ser in copular sentences.”

74
There are certain contexts, however, where Gumiel-Molina et al’s analysis does not seem to be able to account for the data in Spanish. One such context is predicates that involve non-gradable adjectives. The authors’ proposal is that this type of adjective only co-occurs with ser. This does not seem to be the case, as the examples in (72) show:

(72) Alex está {vivo/muerto}.
    Alex ESTAR {alive/dead}
    ‘Alex is {alive/dead}’

Gumiel-Molina et al consider that “since adjectives like these appear in comparative constructions and can be modified by proportional modifiers like medio ‘half’, we consider them gradable”. This analysis, however, seems to focus on metaphorical, or idiomatic uses of these adjectives. The examples in (73) show how both muerto ‘dead’ and vivo ‘alive’ behave like gradable adjectives:

(73) a. Esta planta está más muerta que aquella.
    this plant ESTAR more dead than that
    ‘This plant is more dead than that one.’

    b. Esta planta está medio muerta.
    this plant ESTAR half dead
    ‘This plant is half-dead.’
While it is true that the adjectives in the previous examples accept comparative structures (73a) and proportional modifiers (73b), the meaning that they provide in these sentences is not the literal one. In these cases, the adjective \textit{muerto} ‘dead’ does not mean ‘no longer alive’, but rather ‘in a deteriorated state’. We can see similar distributions of the adjective with the colloquial meaning ‘very tired’ in (74):

\begin{itemize}
  \item[(74) \ a.] Después de ir al gimnasio, estoy medio muerto.
      \hspace{1cm} After \textit{go} to-the gym, \textit{I am} \textit{half} dead
      \hspace{1cm} ‘After going to the gym, I’m really tired.’
\end{itemize}

If, on the other hand, the literal interpretation of \textit{muerto} ‘dead’ is operative, neither comparative structures nor modifiers are acceptable, as can be seen in (75):

\begin{itemize}
  \item[(75) \ a.] #Napoleón está más muerto que Julio César.
      \hspace{1cm} Napoleon \textit{is} \textit{more} \textit{dead} \textit{than} Julius Caesar
      \hspace{1cm} ‘Napoleon is more dead than Julius Caesar.’
  \item[(75) \ b.] #Pepe está medio muerto.
      \hspace{1cm} Pepe \textit{is} \textit{half} \textit{dead}
      \hspace{1cm} ‘Pepe is half-dead.’
\end{itemize}

Another context in which Gumiel-Molina \textit{et al}’s theory seems to falter is in cases where an open-scale adjective can co-occur with both \textit{ser} and \textit{estar}. The authors’ proposal in these cases is that
there is a coercion mechanism, by which “the combination of some adjectives with estar will be more marked in those cases where the adjective expresses a property that is least likely to be conceived as predicated of stages of the subject.” However, once again the problem of rico ‘famoso’ and famoso ‘famous’ resurfaces. These adjectives, as previously shown, pass all the tests for them to be considered Stage Level predicates, and yet, they combine exclusively with ser in Spanish. The test provided by Gumiel-Molina et al. to identify open-scale adjectives that usually combine with ser but can also co-occur with estar is based on identifying them as dispositional adjectives (Krifka et al. 1995; see also Dixon, 1982; Bierwisch, 1989; Stowell, 1991; Arche, 2006, 2011). Dispositional adjectives show certain properties, for example dynamicity—they are possible in the progressive, (76)—as well as agentivity—they are possible in the imperative, (77). These tests, however, also obtain in the case of adjectives like rico ‘rich’, famoso ‘famous’, importante ‘important’ or necesario ‘necessary’:

(76) Lleva siendo rico y famoso desde que se casó.

(He) carries being rich and famous since (he) married

‘He has been rich and famous since he married.’

(77) Sé necesario en tu empresa.

be necessary in your company

‘Be necessary in your company.’
In addition, in those cases where open-scale adjectives can appear in an estar predication, Gumiel-Molina et al.’s requirement that these adjectives be evaluated from the perspective of a within-individual comparison class does not reflect the interpretation of the sentence. The example in (78) show predications involving adjectives that describe emotional states which can all be used with estar:

(78) María está {nerviosa/feliz/triste}.
    María isESTAR {nervous/happy/sad}.
    ‘María is {nervous/happy/sad}’.

The interpretation of (78), according to a strict within-individual comparison class, as proposed by the Gumiel-Molina et al., is that the degree of nervousness, happiness or sadness of the individual María will only be compared to other degrees projected by María on the same scale. This does not, however, generate the right interpretation, since example (78) could be applied to María even if we have never met her.

It seems then that the notion that dispositional adjectives can be systematically coerced into estar predications does not reflect all the data that can be found in Spanish. In cases such as those found in (78), there needs to be a standard degree that needs to be met or exceeded by the individual, and this degree is constructed by comparing relevant individuals, not simply different stages of the same individual. In other words, we need access to a between-individuals comparison class. For example, if we consider the sentence in (79):
(79) María está nerviosa.

María is\textit{estar} nervous

‘María is nervous’

The degree of nervousness of an individual is not simply determined by her own experiences. A Buddhist monk who has lived all her life in a monastery and a member of a bomb disposal squad in the military will have different standards for nervousness. In other words, different standards could be expected for different individuals, and in some cases, vastly different ones. However, in our previous context, if María is a Buddhist monk, a sentence such as (79), still implies the same behavior or condition if María is a normal person. Furthermore, it is possible to produce sentences with overt, between-individual comparison classes with this type of open scale adjectives, something which is not possible with closed-scale ones, as the examples in (80) illustrate:

(80) a. María está nerviosa para un monje budista.

María is\textit{estar} nervous for a monk Buddhist

‘María is nervous for a Buddhist monk’

b. #La puerta está cerrada para una puerta automática.

The door is\textit{estar} closed for a door automatic

‘The door is closed for an automatic door’
Evidential uses of *estar* are another context in which the Gumiel-Molina et al.’s approach struggles to capture the interpretation of the data. According to the authors “in the evidential reading, counterparts are related not to the subject, as just explained, but rather to an implicit experiencer introduced by the evaluative adjective”. In other words, the comparison class is constructed around the experiencer’s opinions. Crucially, this type of comparison class cannot work as a within-individual. This type of comparison class, as outlined in Toledo & Sassoon (2011), is constructed of world-time pairs where the subject of the predication projects different degrees in the scale provided by the adjective. For example, consider (81):

(81)  
La puerta está abierta.  
The door is *open*.  
‘The door is open’.

In (81), the comparison class includes worlds in which the door has different degrees of openness. Given that *abierta* ‘open’ has a lower bound on its scale—when the door is closed—if the sentence expresses any degree other than that lower bound, the door being open will be a true statement. In the case of the evidential uses of *estar*, this type of evaluation is impossible. For example, consider the following sentence in (82):

(82)  
Esta canción está genial.  
This song is *genius*  
‘This song is great’
In (82), trying to construct a comparison class just with different situations in which this particular song was in a different state of excellence makes little sense. Instead, the interpretation implies a comparison between this particular song and others, perhaps of lesser quality. In other words, evidential uses of estar seem to require a between-individuals comparison class, contra Gumiel-Molina et al.’s claim that all estar predicates must include a within-individual comparison class.

2.6. Conclusions
This chapter has provided an insight into the main theoretical approaches that have been proposed to capture the ser and estar distinction in Spanish. By reviewing these influential theories, certain patterns can be identified. Perhaps one of the hindrances to formulating a theory that accounts for copula distribution in Spanish is that certain premises have been taken as an a priori truth—the temporal nature of the distinction between ser and estar—where perhaps a different, deeper analysis is warranted, in view of the data present in the language. As it has been pointed out several times in this chapter, neither ser nor estar carries with it a necessary interpretation of permanency or temporariness, whether it is expressed in terms of SL and IL predicates or aspectual features. It cannot be surprising that the evidential uses of estar present an uphill battle for such frameworks, given their apparent mapping with what have been traditionally taken to be ser contexts. In order to try to transcend these narrow views, there are several essential observations that can be gathered from the extensive research on the topic.
On the one hand, temporal implications in the interpretation of *ser* and *está* predications are a consequence of their distribution, not the driving force behind the distribution itself. As it has been previously mentioned, neither copula carries with it a strict requirement that their predicates denote either a permanent property in the case of *ser*, or a temporary one in the case of *está*. The examples in (83) show this lack of temporal interpretation:

(83)  

a. Las casas rojas están vacías.  
The houses red are*está* empty.  
‘The red houses are empty’

b. Alex es profesor.  
Alex is*ser* teacher  
‘Alex is a teacher’

In (83a), the state of occupancy of the red houses, past, present or future, plays no role in the interpretation of the sentence. The sentence can be true whether the red houses have never been and never will be occupied or not. The same temporal interpretation applies to (83b). The only thing that can be predicated about Alex is that he is a teacher now, but whether he will remain so indefinitely is not specified. Romeu (2015) provides the closest analysis in terms of temporal interpretation, by stipulating a predication with *está* provides access to a different state than the one referred to by the adjective. But, it does not need to be a state that the subject has, or will, posses. Accounts that rely on concepts such as SL oraspectual considerations cannot handle this type of subtle difference, and they require that the subject change the state denoted by the
predication, which leads to wrong interpretation in some cases. In this manner, temporary, cases where SL/IL or aspectual features—in (83), those cases where the red houses were previously occupied, or Alex never changed careers—are present become a subset of all cases of copular predications, not the entirety of the phenomenon.

A second factor that needs to be considered is that there is a type of use of estar that does not seem to follow the same criteria as what has traditionally been considered predication of temporary characteristics of individuals. These are the uses of estar that refer to an evidential use of the copula. These contexts are particularly problematic for theoretical frameworks that rely on the temporal properties of the predications to account for the distribution of ser and estar. These evidential uses of estar frequently involve predications where the subject cannot possibly change the state denoted by the predicate. For example, consider (84):

(84) Esta película está muy chida.

This movie is estar very cool.

‘This movie is very cool’

In cases such as (84), not even the accessibility solution proposed by Romeu (2015) will suffice, since it is not logical to think that a particular movie may become a better or worse version of itself. These evidential uses require on the part of any theoretical account that they further move away from temporal considerations, if they do not wish to give evidential uses of estar an exceptional status within the system.
A theory of copula distribution in Spanish needs to make, as much as possible, the right predictions about which predicates will appear with *ser* or *estar*. While it is true that *ser* and *estar* cover a vast number of contexts, and dialectal variation is well documented in copula choice, certain patterns in their distribution should not be ignored. Many of the theories presented in this chapter succeed in explaining structures present in Spanish, but in doing so make the wrong predictions about other structures that should exist in Spanish, but do not. This is evident in cases such as adjectives as *rico* ‘rich’ and *famoso* ‘famous’. Copular complements that include such adjectives are predicted by many theories to be able to combine with both *ser* and *estar*, and yet they are strict *ser* predicates. Whether we subscribe to SL/IL, aspectual or more classical permanent vs temporary schools of thought, it seems clear that *rich* and *famous* should be able to denote both available options—there are permanently rich and famous people, as well as temporary ones. And yet, this type of adjective is not productive in *estar* predications. In this sense, these adjectives present a challenge not in the sense that their occurrence with *ser* is not explained, but rather because their absence from *estar* predications is not. These gaps in the copular system are systematically left out of theoretical accounts, but nevertheless need to be addressed if the goal is to present an all-encompassing theory of the distribution of *ser* and *estar* in Spanish. Perhaps the fluid nature language combined with the vast array of possibilities that adjectival predicates present cannot be fully accounted for, but that goal should at least be the target a theoretical proposal trying to account for the phenomenon.
In the next chapter, these three considerations that can be distilled from the vast existing literature will be incorporated into a theoretical framework that attempts to give the distribution of both copula choices in Spanish a single motivation.
Chapter 3: Scalar Account for Ser and Estar

3.1 Introduction

At this point, all the different uses of ser and estar in Spanish and an overview of all the different approaches in the literature to account for their distribution and interpretation have been presented. The main objective of the present chapter is to develop a new theoretical proposal that incorporates the ideas from previous authors, as discussed in Chapter 2, and that explains all the different uses of ser and estar described in Chapter 1. In this chapter, I provide a formal account of the distribution of ser and estar within a scalar theoretical framework. I incorporate the basic ideas derived from the observation of the data from Chapter 1, as well as from the previous attempts at capturing the differences in the copulas in Spanish seen in Chapter 2. This chapter first establishes three basic observations that need to be incorporated into the framework, namely the idea that ser and estar are not dependent on considerations of different stages that an individual has experienced, but rather on considerations of potential states, not necessarily attained. The second idea that I incorporate into the framework is the distributions of the copulas being sensitive to the scale structure of the predicates they combine with, more specifically, the presence or absence of logical endpoints on their scale and how adjectives that lack these need a natural process in order to appear in estar predications. The final idea to be incorporated are the evidential uses of estar, and how their peculiarities can be accommodated into the general paradigm of estar predicates. This general paradigm, formulated under the general framework of scalar accounts of adjectives, can be summarized in one overarching statement: predicates that combine with ser are characterized by denoting a single degree on their scale, whereas estar
predicates denote an interval on their scale. This last sentence is the distilled essence around which the rest of the proposal gravitates. As far as the difference between evidential and non-evidential uses of estar, it is only in the way that the interval is generated that they differ, but the necessity of said scalar interval tied both cases together.

Before anything else, there is the need for a caveat. The proposal presented in this chapter centers around the distribution of copulas with adjectival complements. The reason for choosing this context exclusively is based on the fact that this is the only context where the same complement can potentially appear with either copula. Closer analyses of the cases where the choice of copula modifies the interpretation of the proposition should attest my theoretical approach. Although the theoretical framework presented in Chapter 3 presents a promising platform from where non-adjectival contexts for ser and estar could be accounted for, this represents a future avenue for research.

In the previous chapters, the distribution of both ser and estar has been presented from a variety of perspectives. The intention was to obtain a perspective with the widest angle of observation possible, in order to see if any patterns could be ascertained. From this vantage point, several conclusions can be reached. One of the more salient ones is that previous theoretical frameworks presented struggle to account for the entire phenomenon of copula distribution in Spanish—they provide valid accounts for some of the contexts, but there are examples that seem to escape the overall criterion presented. On the other hand, there were several observations that can lead to a better capture of the problem—i) temporal considerations such as permanent vs temporary or SL
vs IL predications are not the driving force behind copula distribution in Spanish, but rather a consequence ii) copulas are sensitive to the scalar structure of their predicates iii) evidential uses of _estar_ need to be carefully considered as they are a source of problematic cases. This chapter attempts to integrate these observations into a system that provides a better insight into the _ser_ and _estar_ dichotomy.

This formal proposal will be formulated in terms of scales and degrees. The reasoning for this choice arises from one of the observations that were derived from the data—_ser_ and _estar_ are sensitive to the scalar structure of their predicates. This was directly derived from examples such as the one in (85):

(85) a. Pinté mi bicicleta. Ahora {es/está} pintada de rojo.

   painted.1st.SG my bicycle now {isSER/isESTAR} painted of red

   ‘I painted my bicycle. Now, it is painted red’

b. Pinté mi bicicleta. Ahora {es/*está} roja.

   painted.1st.SG my bicycle now {isSER/isESTAR} red

   ‘I painted my bicycle. Now, it is red’

---

2 Native speakers of Spanish consulted have expressed different views on the status of sentences such as 1b) with _estar_. I use the ungrammatical notation for these sentences on the basis that corpus searches have failed to provide examples of them. However, I acknowledge the fact that some native speakers of Spanish seem to view them not as ungrammatical, but merely as odd or infrequent. Nevertheless, these differences can be accounted for by noting dialectal variation within the extent of the natural processes explained later in the chapter, and do not represent uses of _estar_ falling outside of the scalar interval requirement. Furthermore, in the case of example (85), the stark contrast between the obligatory use of _estar_ in (85a) and the acceptability of _ser_ in (85b) displays the scalar sensitivity that I wish to highlight, regardless of the varying degrees to which _estar_ is acceptable in (85b).
The two sentences above describe the same situation—the newly acquired red color of the bicycle—and yet, a simple switching of the predication from a closed-scale realization with *pintada* ‘painted’ to an open-scale one with *roja* ‘red’, generates a necessary change in the copula. This thus-far unaccounted-for observation leads to the necessity of selecting a theoretical framework that accounts for the scalar structure of the predications. In the following sections, this distribution will be more closely examined and integrated into the proposed framework.

The second observation that was distilled from the data is that *ser* and *estar*, despite many claims, do not base their distribution on the temporal properties of their predications. For example, as previously noted, a sentence such as *Esa casa está vacía* ‘That house is *estar* empty’ can describe a house that we know has been or will be occupied, but also a house that has never been occupied and we have no knowledge if it ever will be. While it is true that a great many cases of uses of *ser* determine a permanent or inherent property of an individual, and *estar* a temporary or transient one, this can be considered a consequence of their distribution, and not a necessity. This idea is introduced by Romeu (2015), as he considers *estar* predicates as connecting the current state of an individual indicating with another, potential state, albeit not one that was or will be necessarily attained. In other words, the mutable nature of *estar* predications does not need to be captured in terms of a change that has happened or will happen, as a classical temporary or SL theory suggests, but rather as one that could happen. This potential state is based on what the speaker judges to be possible, and in this manner, serves as a bridge between the classic SL/IL, Aspectual theories (Kratzer, 1995; Camacho 2012), which were based
purely on the temporal structure of the predicate, and the pragmatic approaches (Maienborn, 2005), which were based on the speaker’s perception of the predication. A scalar account can provide the necessary tools to describe a potential change in the nature of a predicate, without needing to make use of a temporal coordinate.

The first task, then, is to determine how these intuitions about *ser* and *estar* can be incorporated into a scalar account. The first claim deals with the non-temporal nature of the distinction between *ser* and *estar*, and how it should be viewed as a consideration between the actual state and potential ones of an individual. The term potential in this context needs to be interpreted as a state that could be achieved, but not one that necessarily will be so. In this sense, cases of SL predicates will be a sub-set of all cases of *estar*, and not the total sum of them, as previous theories suggested. The main theoretical claim in the present proposal is that *ser* combines with predicates that denote a single point on a scale, whereas *estar* combines with predicates that determine an interval on that scale. This scalar implementation allows us to account for both the present state of an individual—thus covering the *ser* cases—and, since scales need not be determined in time, can also account for potential states of said individual—thus covering the *estar* cases.

For example, in a classic example of an adjective that can occur both with *ser* and *estar*, *Carlos es aburrido* ‘Carlos is *ser* boring’, the sentence describes a situation in which we are only aware of Carlos’ degree of boredom right now, i.e. a single degree of boredom. On the other hand, *Carlos está aburrido* ‘Carlos is *estar* bored’, describes a situation where Carlos is bored now, and
we are aware that Carlos could potentially not be bored, i.e. we have access to a different degree that Carlos could project in the boredom scale. Section 3.5 in this chapter articulates these notions within the scalar architecture.

Regarding the second claim, *ser* and *estar* are crucially sensitive to the presence or absence of logical boundaries in the scalar structure of adjectives. This distinction is articulated in the literature around the concept of open and closed-scale adjectives (Kennedy & McNally, 2005). Section 3.6 in this chapter deals both with the description of open and closed scale adjectives, as well as how this distinction can be used to account for the distribution of copulas in Spanish.

The third claim, regarding the importance of considering the evidential uses of *estar* as an integral part of the *ser* and *estar* phenomenon in Spanish. Evidential uses of *estar* present an interesting problem for traditional accounts based around temporality or aspectual features of predicates, since they do not seem to follow the general rule (Mangialavori, 2013a; Camacho, 2015). For instance, an evidential use of *estar* such as the one in example (86):

(86) Este jamón está estupendo.

this ham *estar* stupendous

‘This ham is fantastic’

The sentence in (86) escapes the traditional explanations, since a specific ham is normally not fantastic in a temporary way—it either is or it isn’t. In scalar terms, this is articulated around the
concept of the standard of comparison. In order to determine whether this ham is fantastic, we
determine, based on our experience with other hams, whether this particular one meets that
standard or not. In other words, the truth conditions of example (86) are based on the speaker’s
impression of how fantastic this ham is as compared to other hams, and not on the way this
particular ham might taste at a different point in time, potential or not. And yet, these evidential
uses of *estar* are commonplace in Spanish. Any account that attempts at capturing the
distribution of copulas in Spanish needs to have a way to account for examples such (86).
Section 3.11 in this chapter covers these cases of *estar* and incorporates them into the general
theory presented throughout the chapter.

In the following sections, these three claims will be given a formal frame and will be integrated
into the existing scalar approach in general. And, although a wide variety of syntactic contexts
have been examined regarding their copula distribution, the present work will focus, as already
mentioned, exclusively on adjectival predications. This type of predication is the one where the
*ser* and *estar* distinction becomes more evident, given that there are ample situations and
adjectival expressions where both copulas can be used. Following the tradition ranging back
from Bello (1891) to current attempts (Gumiel-Molina *et al.*, 2015), the distribution of copulas
based on adjectival properties will be used as a foundation from where other contexts can be
derived in future investigations.
3.2 Copulas and the interpretation of *ser* and *estar*.

3.2.1 The Verb Phrase and the Predicate Phrase

Before we delve into the scalar considerations of adjectives, we need to define in a formal way the actual interpretation of *ser* and *estar*. One of the crucial observations that has been made thus far is that copula distribution hinges on the nature of their predicates—more precisely, there is a particular property in those predicates that allows them to combine with either copula. In this section, I will describe the three main structures that come into play in the derivation of copulas in Spanish, a Copular Verb Phrase $V_{\text{Cop}}P$, a Predicate Phrase $\text{PredP}$ and a Degree Phrase $\text{DegP}$.

Regarding the first of those structures, following Gallego & Uriagereka (2009), Gumiel-Molina & Pérez-Jiménez (2012), and Gumiel-Molina *et al.* (2015), it is assumed that Spanish has one empty copular verb, $V_{\text{Cop}}$, and that *ser* and *estar* are nothing more than a spell-out reflex of the nature of their predicates. Under this approach, the Spanish copula $V_{\text{Cop}}$ is taken to be a verbalizer (Schmitt, 2005) that allows the set of events introduced by the predication to be accessed by the aspectual and tense operators. Spanish needs this verbalizer mediation since it is a language that cannot establish a connection between a predication and the aspectual and tense phrases directly (Adger & Ramchand, 2003; Rothstein, 2001). The structure of the $V_{\text{Cop}}P$ can be seen in example (87):
The $V_{\text{Cop}P}$ takes as one of its arguments the second of the structures under consideration in this section, the Predicate Phrase. This PredP contains the adjectival expression (Gumiel-Molina et al., 2015) which determines the properties that drive copula choice. The Predicate Phrase appears as the complement of the copula *per se*, a consideration along the lines of those suggested by Bowers (1993), Baker (2003) and Mikkelson (2005). According to these authors, Predicate Phrases are not the exclusive complement of copular constructions, but are also present in secondary predications, as seen in example (88):

\[(88) \quad \begin{align*}
\text{a.} & \quad \text{Vi} \quad \text{a José} \quad \text{preparado}. \\
& \quad \text{saw.1.SG. \quad to Jose \quad ready} \\
& \quad \text{‘I saw José ready’}
\end{align*} \]

\[(88) \quad \begin{align*}
\text{b.} & \quad \text{Considero} \quad \text{a Juana} \quad \text{inteligente}. \\
& \quad \text{considered.1.SG. \quad to Juana \quad intelligent} \\
& \quad \text{‘I consider Juana intelligent’}
\end{align*} \]

It has also been suggested (Cinque, 2010; Demonte, 1999, 2008) that these Predication Phrases are found in post-nominal adjectives in languages such as Spanish, as in example (89):
Gallego & Uriagereka (2009, 2016), and Gumiel et al. (2015) establish, in view of this distribution of the Predicate Phrase, that the copula in Spanish, whether it is realized as *ser* or *estar*, takes a PredP as its complement. This PredP introduces, based on Brownlow (2011), both the subject of the predication—the individual of whom a property is predicated—as well as the property itself—in our case, in the form of a Degree Phrase. This interpretation has been formalized (Gumiel et al., 2015) in the following terms:
In this example, the individual *Alberto* combines with a Degree Phrase, *alto*, to form a PredP. Under the current proposal, if the DegP *alto* determines a single degree on the height scale, this PredP will combine with *ser*, whereas if the DegP determines an interval on that scale, it will combine with *estar*. As mentioned, this DegP that is the Complement of the PredP is where the semantic properties that will determine whether *ser* or *estar* will be realized in the $V_{Cop,P}$ are determined. As such, it is the structure that I will pay more attention to, as it is the crux of the *ser* and *estar* distinction.

3.3 The Degree Phrase

3.3.1 The Degree Phrase and Adjectival Properties

In order to define the structure of a DegP, first there is the need to account for its central component. At the core of the DegP lies the positive morpheme (Kennedy, 1997; Husband, 2011). In the case of the present proposal, and in keeping with the standard literature on adjective scales, the positive morpheme *pos* is a piece of covert degree morphology which binds the adjective scale and the degree projected by the individual and the situation. While Kennedy (1997) proposes a single *pos* morpheme to account for all possible DegP, Husband (2011) argues for multiple ones, depending on the scalar structure of the adjective. In the current proposal, I will follow Kennedy and assume that the *pos* morpheme is the same, regardless of scalar structure, but I will posit that, in the case of Spanish, the *pos* morpheme has a different form in the case of predicates that will be able to appear in *ser* predications and those that will do so in *estar*. As such, there are in fact two different *pos* morphemes, a $pos_{SER}$ and a $pos_{ESTAR}$: the former gives rise to DegP that denotes a property involving a single degree on the adjective scale, and
the latter generates DegP that denotes a property involving a scalar interval on the scale. The nature of these two different pos morphemes, one for ser and one for estar, sits at the core of the formal proposal of the present dissertation. Since the nature of this positive morpheme is scalar in nature, the first requirement is to define what scalar degrees are and how they can be implemented.

3.3.2 Adjective Scales

The first concept that needs to be defined in order to describe the distribution of copulas in Spanish is that of an adjectival scale. The first basic assumption is that adjectives are lexically determined for a scale, which they provide to the semantic interpretation of the sentence. Formally, a scale is defined as a set of degrees ordered along a dimension (Kennedy & McNally, 2005). These degrees are interpreted as the atomic elements of the scale, its basic elements. The dimension of the scale is defined by the adjective itself. For example, the degrees that comprise the scale of an adjective like alto are ordered along a dimension of height.

Two important characteristics of scales are defined as follows:

(90) A set S containing at least two degrees along a dimension D with an ordering relation < is a scale iff

a. for all \( d, d' \in S \), \( d < d' \) or \( d' < d \)

(for any two degrees in S, either one is greater than the other)

b. for all \( d, d' \in S \) such that \( d < d' \), there is \( d'' \) such that \( d < d'' < d' \)
(for any two degrees in S, you can find another degree in between)

Naturally, the first notion that needs to be defined is that of a degree. In this case, a degree is an atomic element which represents abstract representation of measurements. They are formalized as points or intervals ordered along a dimension (Klein, 1991).

The first of these two features allows us to consider that there are no two distinct degrees that are equivalent, and that it is always possible to establish a comparison between two degrees. In the degree intervals, the comparison is established in terms of the initial and final points of those intervals. Two intervals will be distinct provided that they have distinct initial and/or final points. This second provision creates a fine-grained structure in the dimension\(^3\).

3.3.3 Standard Degrees and Comparison class.

As seen in the previous section, an adjective provides a scale, which is a set of degrees ordered along a dimension. For example, the scale of the adjective like alto ‘tall’ is comprised of a set of degrees ordered along the height dimension. In order to understand how an individual enters the semantic derivation and interacts with a scale, there are two further concepts that need to be defined—the standard of the comparison and the comparison class.

\(^3\) One reviewer suggests that certain adjectival expressions do not seem to follow the principle expressed in b) in the sense that they lack this fine grained scale—one-legged man, nine-headed hydra. Whether this turns out to be the case or not, for the purposes of the present study, the nature of adjectival scales regarding their graininess has no bearing. As such, all scales are assumed to be fine grained.
As Kennedy and McNally (2005) point out, “a well-known property of predicative uses of gradable adjectives like *tall* and *expensive* is that their interpretations are context-dependent: what counts as tall or expensive may vary from context to context” (pg. 351). In order to account for this variation in the interpretation of what counts as being *tall* or *expensive*, their truth condition needs to be characterized in terms of a standard of comparison (Sapir, 1944; Wheeler, 1972; Bartsch & Vennemann, 1973; McConnell-Ginet, 1973; Kamp, 1975; Lewis, 1986, Klein 1980, 1991; Bierwisch, 1989; Ludlow, 1989; Kennedy 1999; Graff, 2000; Barker, 2002). The standard of comparison, then, is the degree that individuals needs to meet or exceed in order to be considered as possessing the quality defined by the adjective. For example, the following adjectival predications (91a, 92a) can be paraphrased in terms of the standard of comparison (91b, 92b):

(91) a. Alex is intelligent.
    b. Alex’s intelligence exceeds the standard of intelligence for this context.

(92) a. A Lamborghini Aventador is expensive.
    The price of a Lamborghini Aventador exceeds the standard of expensiveness for this context.

The standard itself is also context dependent, and is determined in relation to a comparison class (Klein, 1980; Bale, 2008, 2011; Toledo & Sassoon, 2011). The comparison class of an adjective depends on the individual it is predicated of and can be of two fundamentally different types: a
set comprised of ‘counterparts’ of that individual (Lewis, 1986), or else an extensional category of that individual (Klein, 1980).

According to Toledo & Sassoon (2011), a counterpart-set comparison class gives rise to a within-individual interpretation, in which the adjective’s argument is compared to its counterparts – realizations of that individual at different situations or possible worlds. For example, in the description of a glass as being full or empty, the interpretation is based on different visualizations of that same glass in various states of emptiness—or fullness, as the case may be. Alternatively, an extensional comparison class will include other individuals that are of a similar nature as the individual which is the subject of the adjectival predications. For example, if Alex is described as intelligent, the comparison class will be comprised of other individuals relevant in the comparison of intelligences—human beings, students, family members, etc. This type of comparison class produces a between-individuals interpretation. The distinction between the within-individual and between-individuals comparison classes is crucial in the generation of ser-valid and estar-valid predicates, as it is shown in section 3.8 in this chapter.

One question that remains unanswered at this point is how the standard of the comparison is determined. Some proposals (Bale, 2008; 2010) suggest that, given a comparison class, the standard of comparison is the mean of all the degree values supplied by that comparison class. Kennedy (1997) argues against this conception, providing a counterexample adapted here:
Alex is slightly taller than the average gymnast. #Alex is tall for a gymnast.

From the previous example, Kennedy argues that it is not possible that the standard of comparison is the mean of the height of the gymnasts, and therefore, the operative mechanism for obtaining the standard must be a different one. He proposes an alternative in which the standard is determined by a function that ranges over the members of the comparison class, but instead of calculating the mean value of the relevant degrees, it obtains a standout degree. This standout degree is the relevant degree for that particular context that an individual needs to meet in order to generate a positive reading of the adjective. The most common implementation of this positive reading is to posit an unrealized degree morpheme pos, (Cresswell,1977; von Stechow 1984, Kennedy, 2004) which has the same semantic type as other degree morphology. Specifically, pos should have the denotation as follows:

\[(94) \quad [\text{Deg } \text{pos}] = \lambda g \lambda x. g(x) > d_s\]

where g is the translation of the adjective, which is also interpreted (Kennedy & McNally, 2005; Husband, 2011) as a degree function associated with the adjective, x is the individual subject of the predication, d_s is shorthand for ‘the contextually appropriate standard of comparison, whatever that is’, and > indicates a greater than relationship, based on the ordering of the scale.

Kennedy (2007) then provides a meaning for the pos morpheme, proposing that it introduces a standard in which the individual “stands out” relative to the scale. Naturally, the nature of how
exactly the standout degree is determined is highly contextual and not easy to capture in a formal way. Although there has been extensive work about the nature of comparison classes (see Bale, 2010 for a review) and the role they play in the evaluation of adjectival predications, the exact nature of how the standard of comparison is computed remains a foggy area.

Beyond the arithmetic nature of the standard, Toledo and Sassoon (2011) argue that the determination of standard is based on three factors.

a. The comparison class evoked in the interpretation of that adjective, which determines the degrees on its lexically encoded scale that are relevant for assigning truth conditions.

b. An economy principle (following Kennedy 2007) which dictates that an interpretation relative to a maximum or a minimum endpoint within a comparison class takes precedence over one relative to an arbitrary midpoint. This principle is known as the Interpretive Economy principle and it stipulates the need to maximize the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions.

c. A grammaticalization principle according to which the type of standard that is usually selected for an adjective is encoded as a default convention. In other words, in a given situation, the standard of comparison is determined as a convention between individuals. For example, in a normal, everyday situation, the standard of comparison for a tall person
could be 6’2”. But, in a situation where professional basketball players are being
evaluated, the standard of tallness could be 6’8”, for example. The process by which a
standard is determined is still a contentious area in the literature, but for the purposes of
the present analysis, it is enough to say that the standard of comparison is a convention
among individuals and that it can change depending on the situation.

Given these features, the nature of the different positive morphemes that generate a predicate that
will produce a *ser* structure and an *estor* structure will also have a reflection on the comparison
classes that these structures require.

### 3.4 *Ser* and *Estar* and the *pos* morpheme

#### 3.4.1 *Ser* and the *pos* morpheme

In section 3.4, we examined the nature and structure of scales. In this section, we will see how
the structure of scales is used to represent the properties of adjectival predicates. To do so, I
redefine the nature of the positive *pos* morpheme that sits at the core of the DegP, since this is
the element that regulates the composition and meaning of the structure. I base my analysis on
the blueprint laid out by Kennedy (1997, 2007), and will make the necessary adjustments to
account for the requirements that *ser* and *estar* place in predicates in Spanish. The main
difference between a predicate that can appear in *ser* predictions and one that can appear in *estar*
predications is that *ser* predicates require the presence of a property defined by a single degree,
whereas *estar* predicates require that property under discussion projects an interval on the scale,
i.e. two degrees are implied.
In the case of the positive morpheme present in _ser_ predications, its formalization is similar to the _pos_ morpheme Kennedy (1997) and Kennedy & McNally (2005) suggest for English:

\[
[[\text{pos}]] = \lambda g_{(e,d)} \lambda P_{(e,t)} \lambda x_{(e)} \cdot g(x) \geq \text{standout}
\]

Where \( g \) is a function that maps individuals to degrees on a scale—this is the function that the adjectival expression supplies, for example, and adjective like _alto_ ‘tall’ is interpreted as a function over individuals and degrees that assigns to those individuals a degree on the height scale—, \( P \) is a comparison class function (Toledo & Sassoon; 2011, Gumiel-Molina et al., 2015) which generates a set of degrees. The exact nature of this comparison class function is explained in section 3.7. Finally, \( x \) is the individual subject of the predication. In our previous example with _alto_ ‘tall’, remembering that we are referring to the positive interpretation of the adjective, as it applies to someone who is considered tall, the structure for the DegP would be:

Additionally, the positive morpheme for _ser_-valid predications requires that the degree projected by the individual \( x \) at the time of the event be equal to or higher than the _standout_ degree
(Kennedy, 1997, 2007) of comparison. This standout degree is obtained in the following manner:

\[(96) \quad \text{standout} = M_{\langle c, d \rangle, \langle (d, t), d \rangle} (g)(P) \]

Where \( M \) is a function that ranges over degree functions (supplied by the adjectival expression) and comparison classes (sets of degrees), and provides a single degree which is used as the standard for comparison given that context. Essentially, given an adjectival expression such as \textit{alto} ‘tall’, and a comparison class made of the degrees of tallness projected by a set of individuals—for instance, the degrees of tallness of professional basketball players, ranging from 5’10 to 7’1— the \( M \) function determines the degree that an individual must meet in order to be considered tall within the context of professional basketball—6’9, for instance.

In the case of adjectival predications that are valid \textit{estar} complements, there needs to be a different property that must be met, based on our observations of the interpretations of \textit{estar} expressions. This property is a scalar interval.

\subsection*{3.4.2 Estar and the pos morpheme}

In the case of \textit{estar} predications, the \textit{pos} morpheme needs to determine two different degrees on that scale, as opposed to the single degree required in \textit{ser} predications. This requirement is derived from the necessity to be able to account for the previously mentioned intuitions about said predicates. First, a valid \textit{estar} predicate needs to be able to establish a relationship not with a
different temporal state of the individual subject of the predication i.e. a change in the state, but with one involving a potential change. In terms of a scalar framework, this can be articulated by the necessary interval involving, on the one hand the degree projected by the individual at the time of the event determined by the sentence, and a second degree that the individual could project. Moreover, these two points that define the interval need to be on different sides of the standard of comparison. For example, in the case of *estar aburrido* ‘to be bored’, the two points would be, on the one hand, the degree that is projected by the individual at the time of the event, which would be a positive interpretation of the adjective and therefore a degree that surpasses the standard for being bored in that situation. On the other hand, the second degree of the interval is one that the individual could project that also falls short of the standard of comparison or, in other words, a degree where the individual would not be bored. This can be represented graphically, as in example (97):

(97)

In example (97), for *estar aburrido*, the individual projects a degree of boredom $d_1$ on the scale of boredom $S$ at the time of the event depicted by the sentence. This degree is bigger than the
standard of comparison for this situation, represented by the standout degree, so the individual receives a positive reading of the adjective—he or she is bored. In order to generate the interval necessary for an estar predication, a second degree $d_2$ needs to be accessible, and, moreover, $d_2$ needs to fall on the other side of the scale as defined by the standout degree. The necessity of this is based on one of the basic premises established in this chapter—estar predications establish a connection between the current state of an individual and a different, potential state. In the case of a scalar implementation, two degrees on the same scale that describe different states requires that they fall on different sides of the standout. In example (97), the requirement is that one degree describes a state of being bored, and the other a state of not being bored. In this manner, we have translated Romeu’s (2015) intuition about estar lexicalizing Disjoint states into a scalar framework.

In order to formalize how the two points on the adjectival scale that are required for an estar predication can be fit into the structure of a DegP, the present proposal will make use of one formal component previously used in the literature in order to solve such a problem: difference functions (Baglini, 2012). According to Kennedy & McNally (1999), gradable adjectives encode measure functions. Measure functions are functions from times to functions from individuals to a degree on a scale defined by the lexical content of the adjective. In other words, what the adjective provides, apart from a scale and its structure, is a measuring function, that maps individuals onto degrees. Difference functions, according to Baglini, are “just like measure functions except that they return a degree representing the difference between an object’s
projection on the relevant scale and an arbitrary degree d, which represents the comparative standard (Baglini, 2012).” She formalizes this concept as can be seen in example (98):

(98) Difference functions: For any measure function m from objects and times to degrees on a scale S, and for any \( d \in S \), \( m_d \) is a function just like m except that:

a. its range is \( \{ d' \in S | d \geq d' \} \) and

b. for any \( x, t \) in the domain of m, if \( m(x)(t) \geq d \) then \( m_d(x)(t) = d \)

Difference functions, then, provide the distance that two degrees project on a scale. On our particular case, the time parameter that both authors include in their formalization can be omitted from the present analysis. Both Kennedy & McNally and Baglini employ measure and difference functions in the study of the interaction between verbal event structure, degree achievements, and adjectival passives, extending the analysis that Dowty (1991) proposed and Kennedy & Levin (2008) later on expanded, that assumes a homomorphic relationship between properties of verbs and the temporal progress of the events they denote in order to account for their telicity. With verbs of change of state, the homomorphism is between a gradable property corresponding to the state associated with the verb; with verbs of motion, the homomorphism is defined by the path traversed by one of its arguments. In the case of defining the properties of the predicates that can appear in a estar predication, the objective is not to establish a parallel with a verbal event associated with the scale, but rather to simply establish the relationship between an individual and an adjectival scale. In such a context, the time parameter that is essential for Baglini and Kennedy & McNally, becomes moot. In fact, in Kennedy’s (2005) analysis of
adjectival scales, the g function that is associated with the adjectival expression—as previously seen—does not include the time argument, and it is simply a function from degree functions to individuals. Removing this temporal component from the measure function represents one of the focal points of the present study. One of the overarching ideas that have been put forward in the present study is that copula choice of adjectival predicates in Spanish is not based on temporal considerations based on actually achieved states, but rather in terms of potential changes in the individual. That is to say, it is not necessary that the individual subject of the predication has attained both states, simply that it could conceivably attained a different state than the one expressed in the sentence. The example in (16), previously seen and repeated here shows this intuition:

(99) Esa casa roja está vacía.

\[
\text{that house red is} _{\text{ESTAR}} \text{ empty}
\]

‘That red house is empty’

The example in (99) does not carry with it any further implication about the occupancy of the red house beyond its present empty state—the speaker is simply indicating that the house is empty now, but whether it has ever been or will ever be occupied is not determined. In other words, a classical interpretation of the sentence in (99) as being temporary or a Stage Level predicate is too strong an interpretation, since it requires that the houses be occupied at some point. While it is true that potential changes are often realized and become SL predicates, strictly speaking, estar predications need not be realized. As such, SL predicates reprent a sub-set of all possible estar
predicates. In the case of deverbal adjectives, however, these potential changes need to be realized, since they are the result of a verbal event (Kennedy & McNally, 1999). This can be attested by comparing the temporal implicatures of deverbal and lexical counterparts such as vaciado ‘emptied’ and vacío ‘empty’. If the intuitions presented here are correct, a estar predicate that includes vaciado will necessarily be the result of an event of emptying, and as such, will denote a temporary state of the subject of the predication. In the case of vacío, on the other hand, there is no event of emptying and could describe an individual that has always been empty, but could be filled. The examples in (100) and (101) try to cancel these implicatures regarding the temporality of the predicates:

(100) a. La piscina está completamente vaciada.
the pool i_{ESTAR} completely emptied
‘The pool is completely emptied’

b. La piscina está vaciada y nunca ha estado llena.
the pool i_{ESTAR} emptied and never has.3rd.SG been full
‘The pool is emptied and it has never been full.’

(101) a. La piscina está vacía.
the pool i_{ESTAR} empty
‘The pool is empty’

b. La piscina está vacía y nunca ha estado llena.
the pool i_{ESTAR} empty and never has.3rd.SG been full
‘The pool is empty and it has never been full.’

Example (100) shows how the temporality of the predication is a necessary condition, since trying to cancel it results in an unacceptable sentence (100b)—the predicate including a deverbal expression vaciada ‘emptied’ is necessarily the result of an event of emptying the pool (Marín & McNally, 2011), and thus, the state of emptiness has changed. In example (101), on the other hand, the lexical adjective vacía ‘empty’ does not carry any temporal interpretation of change, and there is nothing preventing this change from being negated in (101b). The examples in (100b) and (101b) can be seen as lending support to eliminating, in the case of characterizing the predicates that will appear in ser and estar structures, the temporal argument from the difference function provided by Baglini (2012) for deverbal adjectival expressions. A similar approach was adopted by Kennedy (2005) with the measure function (Kennedy & McNally, 1999) when accounting for lexical adjectives. As such, by eliminating the temporal condition from the interpretation of the predicates, we can subsume both deverbal and lexical adjectives into one paradigm, with the former having a necessary temporal interpretation and the latter being underspecified in that respect.

In addition to the already presented argument against including a temporal argument in the derivation of the difference function, there is also another argument related to the general approach to the distribution of ser and estar. In keeping with one of the general assumptions presented in this study, which is that copula choice in Spanish is not based on achieved temporal considerations, but rather on those of potential changes, it seems more logical to not include a
temporal component in the pos morpheme that lies at the core of the Degree Phrase generating a valid estar predication. Given these premises, the difference function for estar predications can be formulated as follows in (102):

(102) Difference function for estar: For any measure function $g$ from objects to degrees on a scale $S$, and for any $d \in S$, $g_{\text{fin}}\uparrow$ is a function just like $g$ except that:

a. its range is $\{ d' \in S \mid d' \geq \text{standout} \}$, and

b. for any $x$ in the domain of $g$, if $g(x) \geq \text{standout}$ then $g_{\text{fin}}\uparrow(x) \geq g_{\text{standout}}\uparrow(x)$

The estar difference function is similar to the measure $g$ function except that the difference function returns a degree which is the difference between the object’s projection on the relevant scale $g(x)$, and another degree on the scale, $\text{fin}$, which represents the other potential state that the object could project. This $\text{fin}$ point is a degree that is in the set of degrees defined by the comparison class $P$. Furthermore, the distance determined by this function needs to exceed the distance on the scale between the degree projected by the object and the standout degree of comparison $g_{\text{standout}}\uparrow(x)$. This carries with it one implication: the degree that the object projects on the scale $g(x)$ and the final point of the interval $\text{fin}$ need to be on different sides of the standout degree of comparison, that is, they would determine states of an individual with different truth conditions regarding the adjectival expression supplying the scale. The mechanism of the measuring function for estar can be seen graphically in (103):
For example, in a sentence such as *Alberto está alto* ‘Alberto is ESTAR tall’, if the individual Alberto projects a degree of tallness $g(x)$ that would qualify him as tall—$g(\text{alberto}) \geq$ standout—the final point $\text{fin}$ of the difference function needs to be one that would not qualify Alberto as being tall. If this were the case, then the distance measured by the difference function between those two degrees, $g_{\text{fin}}(x)$, is greater than the distance between the standout degree and Alberto’s degree of tallness, $g_{\text{standout}}(x)$, and that ensures that the condition that *Alberto está alto* refers to two degrees of tallness of different states is achieved.

In the case of our requirements for a valid *estar* predication, and taking Baglini’s difference functions as a starting point, the difference function $g^\uparrow$ will be as follows:

\[(104) \quad g^\uparrow_{\text{fin}} \ll (<d, t>, <e, (<d>,<e>)(P)>)_{<d,t>}\]

Where $g$ is the measuring function provided by the adjective, final degree $\text{fin}$ is a degree that is a member of the set of degrees in the comparison class $P$ ($\text{fin} = d \in P$).
Given these considerations, the \textit{pos} morpheme, in the case of a DegP that can appear in \textit{estar} predication, will be formally defined as in (105).

\begin{equation}
[\text{pos}] = \lambda g_{(c,d)} \lambda P_{(d,t)} \lambda x_{(e)} \cdot g(x) \succ \text{standout} \land g \uparrow \text{fin}(x) \succ g \uparrow \text{standout}(x)
\end{equation}

In essence, the \textit{pos} morpheme takes an adjectival degree function of individuals into degrees \( g \), a comparison class \( P \) and an individual \( x \), and states the degree that the individual projects on that scale exceeds the standard degree of comparison on that context and that the scalar distance between the degree that projects the individual \( x \) on the scale and the final point \( \text{fin} \) is greater than the distance between the degree that the individual \( x \) projects and the \textit{standout} degree of comparison. Again, this configuration ensures that the degree that the individual projects generates a positive reading of the adjective—Alberto is a tall person, on our previous example—and that there is an accessible degree in this context such that it is further away than the standard of comparison—Alberto could potentially be non-tall.

Given that these two basic \textit{pos} morphemes that generate the desired interpretations of the adjectival predicates, we must now turn our attention to the different distribution of the \textit{estar} copula in open and closed scale adjective contexts. There is a further type of feature that must be driving the distribution of \textit{estar} in these cases, as the simple stipulation of being able to connect the present degree of the individual with a different one does not suffice to account for the
distribution of the copulas in these contexts. The previous example (85), repeated here illustrates this distribution of *estar*:

(106)  

a. Pinté mi bicicleta. Ahora {*es/está} pintada de rojo.
painted.1st.SG my bicycle now {isSER/isESTAR} painted of red
‘I painted my bicycle. Now, it is painted red.’
b. Pinté mi bicicleta. Ahora {es/*está} roja.
painted.1st.SG my bicycle now {isSER/isESTAR} red
‘I painted my bicycle. Now, it is red.’

In both of the examples in (106), the context is essentially the same. Furthermore, the events that cause the change in the color of the bicycle are also the same. And yet, (106a) is a necessarily *estar* structure, whereas (106b) is a *ser* one. This distribution of the copulas, where one structure demands *estar* while the other does not allow it, is not uncommon in Spanish, but it is only found in cases where the adjectival expression has an open scale, never in the case of closed scales. In addition, in closed scale contexts, if a different state is available, then the predication must appear with an *estar* copula, and *ser* is never an option. This is not the case in *ser* predications with open scale adjectives, which can support both copulas. Consider the examples in (107):

(107)  

a. El vaso {*es/está} lleno.  
the glass {isSER/isESTAR} full
‘The glass is full’
b. En otoño, las hojas son/están rojas. ‘In the fall, leaves are red’

In example (107a), the closed scale adjective lleno ‘closed’, when applied to the individual vaso ‘glass’ provides us with access to the necessary potential non-full state—glasses can be empty. As such, estar is the only copula that is acceptable in describing this situation of a full glass. In example (107b), on the other hand, the open scale adjective rojo ‘red’ also allows for the necessary non-red state of the leaves, and, as such, the copula estar is available. But this situation does not prevent ser from being used. As far as the differences in the interpretation when using ser or estar in (107b), they are subtle, but not unpredicted. It is also interesting to notice the difference between (106b) and (107b). In example (106b), the use of estar generates a sentence that most native speakers of Spanish find ungrammatical, or at the very least odd. These examples reflect one of the innovations that the present study proposes—the importance that certain natural processes have in the acceptability of estar predicates with open-scale adjectives in Spanish. In section 3.7 of this chapter, these natural processes and their relevance are examined in detail.

Furthermore, if the estar version is produced in sentences such as (107b), there is an implication that the leaves change color and turn red in the fall, whereas if the ser copula is used there is no such implication, and all that is being predicated is that leaves are indeed red in the fall, without
any further information being provided. This interpretation can be attested by trying to cancel the suggested change, as in (108):

(108) a. En otoño, las hojas son rojas y nunca cambian.
   in fall the leaves are red and never (they) change
   ‘In the fall, leaves are red and they never change’

b. #En otoño, las hojas están rojas y nunca cambian.
   in fall the leaves are red and never (they) change
   ‘In the fall, leaves are red and they never change’

As can be seen in (108a), cancelling the change does not change the acceptability of the sentence, which is not the case in (108b), which must necessarily imply the change. This type of sensitivity to the scalar structure needs to be accounted for in any theoretical framework that tries to provide an explanation to the distribution of copulas in Spanish.

3.5 Open and Closed Scales and estar

3.5.1 Open and Closed Scales and copula distribution

Kennedy & McNally (2005) and Kennedy (2007) argue that the scales of gradable adjectives come in four different forms: open scales, lower-closed scales, upper-closed scales and totally closed scales. An open scale is one where there is neither a maximal nor minimal point to it. A lower-closed scale has a minimal point i.e. a degree such that there is no degree smaller than it.
An upper-closed scale is one with a maximal point i.e. a degree such that there is no degree bigger than it. A fully-closed scale is one with both a maximal and a minimal degree.

Evidence for this typology can be found in the distribution of certain degree modifiers such as ligéramente ‘slightly’ and perfectamente ‘perfectly’, which pick out, respectively, the minimum and the maximum degree on their adjectival argument’s scale. For example:

(109) Typology of Scale Structure

a. Open (alto ‘tall’) ➞ (ligéramente, perfectamente) alto
b. Lower closed (sucio ‘dirty’) ➞ (ligéramente, perfectamente) sucio
c. Upper closed (limpio ‘clean’) ➞ (ligéramente, perfectamente) sucio
d. Totally closed (lleno ‘full’) ➞ (ligéramente, perfectamente) lleno

The presence or absence of endpoints on a scale is relevant to the ser and estar distribution in Spanish, as the copulas do not behave in the same manner with open and closed scale adjectives. This is a property that has been previously unmentioned in the literature and will be discussed and integrated into the formal account in more detail in following sections. As an introduction to the phenomenon, the following examples can serve to illustrate the importance of the scalar structure of the adjective involved in the copular predication. The example in (110), containing a closed scale adjective, which displays how the capacity for change affects copula choice:
In example (110a), doors can be opened and closed, and as such, the other state—closed and open—is available to us. This produces a situation where, when describing the situation of a door regarding its open or close situation, it is always expressed with estar in Spanish. It is worth noticing that, as (110) reflects, it is not the adjective per se that determines copula choice in Spanish, but rather the context of the sentence, expressed in the comparison class under the theoretical framework presented here.

As far as open scale adjectives, the situation is a bit different. First, let us examine in example (111) a case similar to the one presented in (110) for closed scale adjectives:

(111) a.  La hojas de pino {son/*están} verdes en primavera.

the leaves of pine {areSER/areESTAR} green in spring

‘Pine leaves are green in Spring’

4 Notice that the sentence La puerta es cerrada ‘The door isSER closed’ is grammatical in Spanish, although only as a passive construction. Under the present study, centered exclusively on adjectival complementation, the adjectival predication version of (110b) is the one intended.
In the case of the examples in (111), the situation is very similar to the one that the closed scale adjectives presented in (110). The color of the pine leaves is such that it surpasses the standard for being green, so ser is a valid option. If the subject of the predication cannot change the property described by the adjective—pine trees are always green—then estar results in an odd choice. If, on the other hand, the subject does provide access to a non-green state—fig tree leaves change color in the fall—then the estar prediction is readily available. In the latter case, there is a difference between open and closed scales. While closed scales require that the estar option be used if it is available as in (110a), this is not the case in open scale adjectives, which allow both the ser and estar options, as in (111b). This particularity will be examined and explained in following sections, but for now, it is enough to show that the ability to access a different state of the positive one expressed by the adjective seems to be a requirement in order to generate a valid estar predicate. But, this does not tell the whole story, as there are cases of open scale adjectives that do not follow this pattern. For example, in (112):

(112)  a. Después de ganar la lotería, Juan {*era/estaba} feliz.  
OPEN SCALE

‘After winning the lottery, Juan was happy’
b. Después de ganar la lotería, Juan {era/*estaba} rico.  
After of win the lottery, juan {was_ser/*was_estar} rich

‘After winning the lottery, Juan was rich

It is fairly clear that both the adjectives in (112a) and (112b) determine states that can potentially change for the given individual. In fact, these are strictly temporary ones, as it is expressed that winning the lottery caused Juan’s happiness and richness. It follows that, as previously expressed, both are predicted to be productive estar situations. In the case of contento ‘happy’ the predication is confirmed and estar is a valid option. And yet, in the case of rico ‘rich’, estar is not acceptable, and instead, we find ser. This unexpected behavior of open scale adjectives can be attested in a multitude of cases and is a crucial distinction that needs to be considered when formulating a theoretical proposal of the distribution of ser and estar. At this point, it is worth mentioning that the deciding factor for whether an open or close adjective can combine with estar in non-evidential uses does not lie in the scalar structure per se, but rather in how these adjectives are evaluated. Open and closed-scale adjectives impose different requirements on these comparison classes, but it is not the scalar structure that determines whether they will be able to appear in ser or estar predications. As mentioned, further sections will examine these cases in more detail, but for now, it seems clear that it is necessary to bear in mind the differences in the scalar structure of adjectives vis-à-vis the presence or absence of endpoints in their scale. And yet, in all cases, every single case of a estar predication, regardless of its scalar structure, carries with it the interpretation that the subject of the predication could potentially be conceived as being in a different state than the one described in the sentence—in (112), Juan
could unhappy or poor. It seems clear that this requirement that was imposed on the *pos* morpheme previously proposed for *estar* predicates, while not sufficient, is a necessary condition.

The open/closed scale dichotomy plays an important role in the distribution of *estar* in Spanish. The basic premise—*estar* requiring a predicate that denotes two points on a scale, and *ser* requiring a predicate denoting a single point—is not sufficient to account for the distribution of adjectival complements in Spanish, in view of how it fails to account for the data in (112b). Under the basic premise, a subject projects a degree on an adjectival scale by virtue of taking part in its predication. Furthermore, if that subject can project a degree which does not exceed the *standout* degree of comparison—an analog of the potential state in Romeu (2105)—it should be able to generate the two points necessary, and therefore, a valid *estar* predication. And yet this does not seem to be the case, as the example in (113) shows:

(113) Pinté mi bicicleta. Ahora {es/*está} roja\(^5\).

*painted.1.SG my bicycle now {isSER/isESTAR} red*

‘I painted my bicycle. Now, it is red’

---

\(^5\) Kennedy & McNally (2008) argue for color adjectives being ambiguous between gradable and non-gradable interpretations. The gradable, open scale interpretation refers to hues or shades of a particular color and the non-gradable, closed scale one to surface covered by the color. Unless otherwise noted, all examples of color adjectives used in this dissertation refer to the gradable interpretation of colors.
The fact that *estar* predication is not an option in (113) is surprising, given that, when evaluating the noun *bicicleta* ‘bicycle’ with respect to the property *roja* ‘red’, we should conceivably be able to access the two points on the scale that the *pos* morpheme for *estar* predications require. The positive red interpretation is a given, since the bicycle is red now, and the converse, non-red degree should also be readily available, as the preceding sentence explicitly states that the subject has been painted yesterday, and, logically, had a non-red degree. This type of distribution can only be observed in open scale adjectives (114a), never with closed scale ones (114b):

\[(114)\]

\[\begin{array}{cccc}
\text{a.} & \text{*Pinté la bicicleta.} & \text{Ahora está *roja.} & \text{Open Scale} \\
& \text{painted.1.SG my bicycle} & \text{now *ESTAR red} & \\
& \text{‘I painted my bicycle. Now, it is red’} & \\
\text{b.} & \text{Pinté la bicicleta.} & \text{Ahora está *pintada de rojo.} & \text{Closed Scale} \\
& \text{painted.1.SG my bicycle} & \text{now *ESTAR painted of red} & \\
& \text{‘I painted my bicycle. Now, it is painted red’} & \\
\end{array}\]

The data in (114) is very revealing of the importance of scale structure and *estar* distribution in Spanish. Both examples are set in the same basic context and refer to a similar property of the subject, namely its color. But, while *rojo* ‘red’ is an open scale (the fact that #100% *roja*, #0% *roja* are unacceptable tells us there is neither an upper nor a lower limit to that scale) *pintada* ‘painted’ (100% *pintada*, #0% *pintada*) is a closed scale. And yet, one of them is a valid *estar*
predication, while the other is not. Given this particular distribution, it seems clear that the choice of the predicates that both ser and estar select is sensitive not only to the capacity of the individual to relate to a different scalar truth condition, but also to the scalar structure itself.

Of course, there is the important question of why some open-scale adjectives can generate valid estar predicates where others cannot, and what sets them apart. The answer lies in the presence or absence of a natural process that drives the change in the property defined by the adjective. Open-scale adjectives that can combine with estar have such a process, while those that do not, lack said process. In order to capture this peculiarity in the distribution of estar, it can prove fruitful to observe this phenomenon within the context of color adjectives. Color adjectives are all open scale adjectives that can systematically appear in ser predications and, under certain circumstances, also in estar ones. Given the wealth of contexts where color adjectives can appear, their close examination can provide the basis for the understanding of estar in open-scale adjectival contexts. The conclusion is that open-scale adjectives in estar predications, in addition to describing a scalar interval, also require the presence of a natural process that could make the potential change in color happen. The following section deals in depth with these processes.

3.5.2 The Natural Process

We turn our attention now to the nature of the processes that allow open-scale adjectives to appear with estar. As a first approximation, let’s consider rojo ‘red’ and the contexts where it
can (115) and cannot (116) appear with estar. The following examples provide some of such contexts:

(115)  a.  En otoño, las hojas de los árboles están rojas.
        in fall the leaves of the trees are*ESTAR red
        ‘In the fall, the leaves on the trees are red’

       b.  Cuando tomas mucho el sol, tu piel está roja.
        when you take much the sun your skin is*ESTAR red
        ‘When you take too much sun, your skin is red’

       c.  Tu nariz está roja cuando te resfrías.
        your nose is*ESTAR red when you.REFL cold
        ‘Your nose is red when you get a cold’

       d.  La piel está roja de tanto rascarte.
        the skin is*ESTAR red of so much scratching
        ‘Your skin is red from so much scratching’

(116)  a.  #Después de pintarla, la bicicleta está roja.
        after of paint it the bicycle is*ESTAR red.
        ‘After painting it, the bicycle is red’

       b.  #Hemos teñido la camiseta y ahora está roja.
        have.1st.PL dyed the shirt and now is*ESTAR red
        ‘We have dyed the shirt and now it is red’
The most striking feature of the previous sentences—at least from a traditional perspective—is that all of them are depicting clearly temporary properties of their respective subjects. And yet, only the ones in (115) are productive with *estar*. In order to capture the feature or features that are driving this distribution, one first hypothesis is that all acceptable cases of *estar + color* predicates involve the subject undergoing a natural process of changes. If the examples in (115) are compared with those in (116), it is clear that there is a presence of one such process in the former—leaves changing color, skin getting sunburnt, sore or scratched—that is absent in the latter. It is also worth mentioning how, once again, the scale structure of the adjective is one of the key factors in copula choice in Spanish. The examples in (116) are not valid *estar rojo* contexts, and yet, in the same context, and in describing the same event, if the open-scale adjective *rojo* ‘red’ is switched to the closed-scale *pintada* ‘painted’, then the copula choice changes, and *estar pintada* ‘to be painted’ when describing the bicycle in (116a) becomes a valid structure. If this analysis is expanded to other color contexts, as the examples in (117) show, the same pattern can be found:

(117) a. El ojo está negro por culpa del golpe.
    the eye is\text{\textit{ESTAR}} black for blame of-the blow
    ‘The eye is black because of the blow’

b. Los plátanos viejos están negros.
    the bananas old are\text{\textit{ESTAR}} black
    ‘Old bananas are black’
c. *El barco está negro después de pintarlo.

   the ship is\text{ESTAR} black after of paint it
   ‘The ship is black after painting it’

d. La cara del ahogado estaba azul y morada.

   the face of-the drowned was\text{ESTAR} blue and purple
   ‘The face of the drowned was blue and purple’

e. *He teñido la camiseta y ahora está azul.

   (I) have.1\text{st}.SG the t-shirt and now (it) is\text{ESTAR} blue
   ‘I’ve dyed my t-shirt and now it is blue’

As it was the case in \textit{rojo} ‘red’ contexts, subjects in color predication can participate in \textit{estar} structures in contexts of being the subject of a natural process that changes their color. On top of being cases where the color of the subject can project a degree on the color scale that does not qualify as a positive reading—in other words, on the other side of the standard of comparison—there is an additional requirement for a process driving that change. It is also worth noticing that, in the case of these \textit{estar} structures with open scale adjectival predicates, their reading needs to be one of a temporal nature. Since the natural process is the driving of the change, and it is a necessary condition for the presence of \textit{estar}, it needs to be the case that the subject has changed its color.

Observing the behavior of \textit{estar} with color expressions, we can formulate the hypothesis that open-scale adjectives require not only a connection to a state where the subject does not have the
color property—the eye or the bananas being non-black in (34), for instance—but also a process that drives that change—fruit rotting. If the analysis is extended to other open-scale adjectives, we encounter one of the most intriguing set of adjectives, as far as copula distribution is concerned, as in (118):

(118) a. Tras aparecer en *Trainspotting*, Ewan McGregor {era/*estaba} muy famoso.
   after appear in *Trainspotting*, Ewan McGregor {wasSER/* wasESTAR} very famous
   ‘After appearing in *Trainspotting*, Ewan McGregor was very famous’

b. Los ganadores de la lotería {son/*están} ricos.
   the winners of the lottery { areSER/* areESTAR} rich
   ‘Lottery winners are rich’

c. Pepe {era/*estaba} muy importante después de ser elegido presidente.
   Pepe{wasSER/* wasESTAR} very important after of be elected president
   ‘Pepe was very important after he was elected president’

d. Ahora que tienes un hijo, un seguro médico {es/*está} necesario.
   now that have.3rd.SG a son, an insurance medical {isSER/* isESTAR} necessary
   ‘Now that you have a son, medical insurance is necessary’

These examples are some of the most challenging cases to account for under traditional perspectives. They all describe unambiguously temporary conditions, and yet, they are not productive in *estar* predications. Furthermore, unlike in the case of color adjectives, the
adjectives in (118) do not seem to ever occur in a context where there is an estar-valid process active.

Besides the adjectives already mentioned in (118), the adjectives in (119) also display this behavior regarding their copula distribution include importante ‘important’, trivial ‘trivial’, conocido ‘known’, or sabido ‘known’, absurdo ‘absurd’, lógico ‘logical’ and fiel ‘faithful’, infiel ‘unfaithful’.

(119) a. En la Edad Media, los herreros {eran/*estaban} importantes.
   in the Age Middle the blacksmiths {wereSER/*wereESTAR} important
   ‘In the Middle Ages, blacksmiths were important’

b. La reina Ginebra no {fue/*estuvo} fiel al Rey Arturo.
   the queen Guinevere not {wasSER/*wasESTAR} faithful to the King Arthur
   ‘Queen Guinevere was not faithful to King Arthur’

c. Creer en la Tierra plana {es/*está} absurdo.
   believe in the Earth flat {wasSER/*wasESTAR} absurd
   ‘Believing in a flat Earth is absurd’

d. Vacunar a los niños pequeños {es/*está} lógico.
   vaccinate to the children small {isSER/*isESTAR} logical
   ‘Vaccinating small children is logical’
A mirror image of the above group are adjectives that describe emotional or mental states. These open scale adjectives are always productive with *estar*, and, in fact, they are often mentioned in second language classes as some of the prototypical examples of adjectives that combine with *ser*. This distribution is predicted by the previously mentioned condition: mental states, as in (120), are the result of an individual going through a process that causes it. They are, in this sense, similar to color adjectives—a cold turns noses red, just like a dull movie can lead to boredom.

(120) a. Alex *era/estaba* aburrido.
    Alex *wasSER/wasESTAR* bored
    ‘Alex was a boring person/was always bored’

b. Los niños *son/están* muy felices.
    the children *areSER/areESTAR* very happy
    ‘Children are very happy’

As the examples in (120) show, these adjectives depicting mental states are not only able to appear in *estar* predications, but often also in *ser* ones. As expected, when in an *estar* predication, they denote a state with the potential to change—Alex is bored right now, and the children are amused at the moment—and a condition that does not allow for change in *ser* ones—Alex is a boring person and the children are always happy. There are also adjectives belonging to this group that do not appear in *ser* predications, as they denote conditions that
cannot be permanent, usually because they are the result of a disease or condition, as can be seen in the examples in (121):

(121) a. Alex *era/estaba* deprimido.
    alex *is/SER /ESTAR* depressed
    ‘Alex is depressed’

b. Los niños *son/están* muy excitados.
    the children *are/SER /ESTAR* very excited
    ‘Children are very excited’

In (121a), depression is not a condition that human beings are born with, and, as such, it is always taken to be potentially reversible. Likewise, in (121b), children are not in a state of ever excitement, and, as such, there is always the possibility that they would calm down. This *estar* exclusive distribution is characteristic of adjectives that describe an illness or injury (122), for the same reasons that were just explained:

(122) a. Alex *era/está* resfriado.
    alex *was/SER /wasESTAR* cold
    ‘Alex had a cold’

b. Los niños *son/están* muy doloridos.
    the children *are/SER /areESTAR* very sore
    ‘Children are very sore’
So far, the open-scale adjectives that have been examined seem to corroborate the hypothesis that open-scale adjectives require, on top of the potential of the subject to change their state, a process that guides that change. One final piece of evidence that highlights the importance of this process in copula distribution can be seen in contexts where one of the aforementioned processes stops. For example, let’s consider the sentences in (123):

(123)  a. Alex es alto.
alex is\text{SER} tall
‘Alex is a tall person’

b. Alex está alto.
alex is\text{ESTAR} tall
‘Alex is tall (has grown)’

In (123a), the only information that can be derived is that Alex is a tall individual when compared to the relevant members of the contextual comparison class. In (123b), on the other hand, Alex is not simply a tall individual, but his degree of tallness is related to another, non-tall degree. Since alto ‘tall’ is an open scale adjective, it is predicted that there is a process that is driving Alex’s change in height. In the case of human beings, this process is, naturally, the normal growth process that kids undergo. Of critical importance to the suggested interpretation of estar predicates with open scale adjectives is that this process must be active at the time determined by the copula for the predicate to be valid. In the case of (123b), it follows that Alex
must be currently growing at the time of utterance, the sentence being in the present tense. And therefore, Alex must be a child, since the growth process is absent in adults, and with it the possibility of alto being a valid estar predication. This prediction can be tested by trying to cancel it and observing the results, as in (124).

(124) a.  Alex tiene 40 años y es alto.
alex has 40 years and is\text{ser} tall
‘Alex is 40 year old and is tall’

b.  #Alex tiene 40 años y está alto.
alex has 40 years and is\text{estar} tall
‘Alex is 40 year old and is tall’

As expected, (124a) works even if the possibility of the subject being in the middle of the growth process is cancelled. Ser predications need no such requirement. Likewise, the oddness of (124b) is also expected, since a 40-year-old is expected to be fully grown, and therefore, there is no natural process to validate the estar alto predication. Although we can conceive a human being such as Alex as being tall or short, and therefore, we should be able to construct an estar-valid scalar interval, alto ‘tall’ is an open scale adjective, and it further requires the presence of a natural process that drives the change in tallness, something that a closed-scale adjective would not require.
In other words, if the process driving the scalar change stops, the possibility of appearing in an *estar* predication does so too. It seems clear then that our initial observation of the distribution of *ser* and *estar* with open scale adjectives being contingent on a process holds across the set of open-scale adjectives. Furthermore, one possible *addendum* to the process hypothesis, in view of the behavior of adjectives such as those in the previous examples, is that the process needs to be one that is borne out of the individual itself. All of the valid *estar* examples of a process seen so far, involve a change that is, at least partially, coming from the individual—a nose turning red, a kid being bored—whereas those adjectives that lack the process seem to be more centered on changes not generated by the individual himself—someone becoming rich, or important.

Although the nature of the *estar*-validating process bears further examination, for the purposes of the present study, it is enough to identify its importance in the distribution of copulas in Spanish.

In order to have a formal way of including the open vs close scale distinction into the scalar system, several authors (Bale, 2008, 2011; Toledo & Sassoon, 2011, among others) have proposed that their respective comparison classes can provide the necessary tools. Comparison classes, as previously seen, are comprised of the necessary individuals to generate the standard of comparison to evaluate the truth condition of the positive form of an adjective. Formally, the distinction between open and closed scales is determined by the comparison class: open scale adjectives require a between-individuals comparison class, whereas closed scale adjectives require a within-individuals comparison.
3.6 Comparison classes

3.6.1 Between-individuals comparison class.

In section 3.5, we saw how the scale structure of an adjective affects its capacity to appear in
*estar* construction in Spanish: closed-scale adjectives (such as vacío ‘empty’, abierto ‘open’) require a scalar interval, whereas open-scale adjectives (such as inteligente ‘intelligent’, divertido ‘fun’), additionally require a natural process to combine with *estar*. In section 3.6, discussed the importance of this interval that open-scale adjectives in *estar* predications require and how its presence is a necessity. In this section, I will integrate those notions into a scalar framework, by means of the difference in the comparison classes between open and closed-scale adjectives. The first task is to define what is a comparison class. A comparison class is a set of degrees that determines the standard of comparison for assessing if a particular individual possess the quality specified by an adjective. For example, in a sentence like *Alberto is tall*, the comparison class is the set of degrees of other relevant people that we use to determine the degree of height that Alberto must have in order to be considered tall.

In the case of open-scale adjectives, their correct interpretation requires a between-individuals comparison class. That is to say, the way in which we interpret these adjectives is based on comparing the degree of the subject of a predication with the degrees of a group of individuals, all of which possess the quality described by the adjective to varying degrees.

The reason why open-scale adjectives require a between-individual comparison class is that they lack endpoints on their scale, and this lack of contextually-independent points requires a standard
of comparison that is computed based on the degrees obtained from other individuals of the same kind as the subject of the predication. For example, the interpretation of an open-scale adjective like tall hinges on comparing the individual subject of the predication with other individuals. If we want to check whether the individual Alberto possesses the quality of being tall, we must compare his height with that of other individuals.

Open-scale adjectives is the basic case of the scalar structure of the adjective defined by Kennedy (2011) when providing his description of the scalar structure of adjectives. Toledo and Sassoon (2011) extend the analysis, and describe this type of comparison class, where Gumiel-Molina et al (2015) extend it to apply to their interpretation of ser and estar.

The comparison class of an open-scaled adjective can now be constructed. For example, given the following predicate involving an open-scale predicate, as in example (125):

(125)  El niño es inteligente.

the boy is_{ser} intelligent

‘The boy is intelligent’
Following Gumiel et al. (2015), I assume the following structure for the comparison class.

\[
\begin{array}{c}
PP \\
P_{(d,t)} \\
\quad C \langle e, (d,t) \rangle \\
\lambda y\langle e \rangle. P(y) \\
(\text{for}) \\
\quad \text{a basketball player}
\end{array}
\]

The comparison class is normally instantiated by a Preposition Phrase headed by *for* in English or *para* in Spanish (Ludlow, 1989; Contreras, 1993), as in *Alberto is tall for a basketball player*. According to Gumiel *et al.*, this PP is in fact interpreted as a property and the preposition—*for* in English or *para* in Spanish—does not receive a semantic interpretation. When the comparison class is not phonetically realized, it is assumed to be realized as a null pronoun C (Stanley, 2000; Kennedy, 2007).

In the case of example (125), the first step is to determine which kind does the individual subject of the predication belong to. Since no comparison class is overtly specified, the default would be the general class of children of this age. Given this, the function P, which generates the set of degrees that comprise the comparison class for that context, will then examine all the individuals contained in that class, and will proceed to generate the appropriate degrees in the present predication—the degree of tallness of all children relevant in this situation. This set of degrees will be the one that will be used to determine the standout degree (Kennedy, 2005, 2007) that will be used as the standard for this particular context.
3.6.2 Within-individuals comparison class.

Just like open-scale adjectives are characterized by not having logical boundaries in their scales, closed-scale adjectives are adjectives that are characterized by having such boundaries. These boundaries affect the type of comparison classes that these adjectives require for their interpretation. In the case of closed-scale adjectives, their comparison classes are not comprised of a set of degrees provided by other individuals, but rather by a set of degrees that the individual subject of the predication could possess across different situations.

In the case of closed scale adjectives, Toledo & Sassoon (2011) propose that the comparison class is “a set comprised of ‘counterparts’ of that individual (Lewis 1986), […] A counterpart-set comparison class gives rise to a within-individual interpretation, in which the adjective’s argument is compared to its counterparts – realizations of that individual in different indices.” I will follow Gumiel et al. (2015) basic formalization of the comparison class for Spanish (127), with one main difference.

(127)
While the basic premise is preserved—it is still necessary to have access to relevant, similar worlds where the relevant properties of the individual hold—the nature of the type of elements the comparison class generates changes. Where in Gumiel-Molina et al. the output of the function is, following Toledo and Sassoon, a set of stages where the properties might express a different realization for the individual subject of the predication, the current proposal is that the comparison class will generate a set of degrees. These degrees will express the different realizations of the relevant property at different situations, but that are ensured to be relevant to the current one expressed in the context of the predication by the accessibility relationship A. In this case, the unrealized C morpheme will take the form in (128):

(128) \( C = \lambda w(s). \forall w'(w'Aw)[(x \text{ is } R(\text{ealized}) \text{ as } d \text{ at } w' \wedge P \text{ at } d \text{ in } w')] \)

In (128), the C morpheme is a function that takes a situation \( w \), and states that for all situations \( w' \), that are accessible from \( w \) by means of the accessibility function \( A \), the individual \( x \) is realized as the degree \( d \) in any situation \( w' \) and the function \( P \) holds for the degree \( d \) in \( w' \). The function \( P \) generates a set of degrees that the individual subject of the predication could have at different stages. This comparison class is therefore within individuals, the individual \( x \) in this case being the argument of the predicate. The accessibility relation \( A \) ensures that only the pertinent situations are evaluated—only those situations where the properties that are necessary for the interpretation of the adjective hold. Furthermore, the function \( P \) will extract those necessary properties. For instance, with an overt comparison class such as in example (129):
In this case, the comparison class will take the situation *para un jueves* ‘for a Thursday’, and will generate all accessible alternative worlds where it holds that the restaurant is at a Thursday point in time and has the properties defined by P, in this case, fullness. The comparison class then will be comprised of all the degrees in accessible stages, i.e. situations, where the restaurant projects a certain degree of fullness. These degrees of fullness of the restaurant at different stages are evaluated in order to determine the *standout* degree of comparison.

The main advantage to this new implementation is that there is some commonality between between-individuals and within-individuals comparison classes: they both involve a set of degrees—and also, the M function (Kennedy, 2005, 2007) that determines the *standout* degree of comparison also has the same structure in both types of comparison classes. For example, returning to the previous sentence:

(130) El restaurante está lleno para un jueves.

(the restaurant is*ESTAR* full for a Thursday)

‘The restaurant is full for a Thursday’
The proposed C morpheme will generate the comparison class by taking the situation \textit{para un jueves} ‘for a Thursday’ and will find all the accessible worlds where it holds that the restaurant has a certain degree, and the function $P$ ensures that the degree is the required by the predication, in this case fullness. Once all the relevant, accessible worlds have been evaluated, the output is the set of occupancy degrees that the restaurant projects on worlds where the normal situation holds (Asher & Morreau, 1995). These are the degrees that the M function will evaluate to generate the \textbf{standout} degree of comparison. The difference function $g^\dagger$ will also evaluate the degrees contained in the comparison class in order to find an appropriate initial point of the \textit{estar} interval.

In view of this mechanism for the implementation of comparison classes, let us now consider the following copular sentence containing a closed scale adjective, \textit{lleno} ‘full’ in (131):

(131) El vaso de Juan está lleno.

\begin{itemize}
  \item the glass of juan is \textit{estar} full
  \item ‘Juan’s glass is full’
\end{itemize}

In this case, since the predication involves a closed-scale adjective, following Toledo & Sassoon (2011), a within-individual comparison class is necessary to correctly interpret the sentence. More precisely, in order to determine the meaning of the glass being full, we need to examine the present state of fullness of the glass with other potential situations in which the same glass might
have had a different degree of fullness. This comparison will determine that the glass is full if there are no other degrees in the comparison class such that the glass being fuller.

In example (131), there is no overt comparison class—no para clause, such as para un domingo ‘for a Sunday’. As such, the default comparison class extends to all possible degrees of fullness of the glass in all situations. In other cases, this situation can be overtly restricted by specifying a set of circumstances. For example, the example in (132) could be modified with an overt situation as follows:

(132) El vaso está lleno para ser un jueves.
    the glass is full for to be a Thursday
    ‘Juan’s glass is full for being a Thursday.’

In this case, the current situation has an overt modifier—it is a Thursday—that would affect the set of accessible worlds, since only those where the day of the week were a Thursday would be considered. In our original example (132), however, the situation is not specified, and thus the statement is evaluated across all possibilities.

The accessibility relation A will provide, given the present world w, the necessary accessible worlds. In the present case, for example, all the worlds where the glass exists, and where it can be filled and emptied, and where the scale of the adjective maintains its scalar structure—worlds where lleno is not a closed scale will not be accessible. Once all these possible, accessible worlds
have been determined, the function $P$ will generate the degrees of fullness that the glass projects at all the different worlds $w'$. This will generate a comparison class comprised of a set of degrees that include all the degrees of fullness that the glass subject of the predication, in the relevant worlds. These are the degrees that will be used to determine the standout degree (Kennedy, 2005). Since the present case involves an upper closed scale, this stand-out degree serves as the maximum point of the scale. In other words, the necessary degree that an individual needs to project in order to be considered full is the maximum degree of fullness.

### 3.6.3 Comparison classes and the pos morpheme.

Now that the different comparison classes have been introduced and formalized, they can be applied to the derivation previously indicated in the pos morpheme. This section describes how a PredP can be completed, and how the structure of PredP accounts for the ser and estar distribution.

More specifically, this section focuses on how the process for constructing an estar predication with open scale adjectives can be formalized. The present account is centered around the presence or absence of the necessary degrees in the set determined by the comparison class. As previously defined, the comparison class in a estar predication needs to provide two degrees along that scale. Example (101), repeated here, defined the nature of these two points as follows:

\[(133)\] **Difference function for estar:** For any measure function $g$ from objects to degrees on a scale $S$, and for any $d \in S$, $g_{\text{fin}}$ is a function just like $g$ except that:
a. its range is \( \{ d' \in S \mid d' \geq \text{standout} \} \), and

b. for any \( x \) in the domain of \( g \), if \( g(x) \geq \text{standout} \) then \( g_{\text{fin}}(x) \geq g_{\text{standout}}(x) \)

The difference function \( g_{\text{fin}} \) determines the interval that the individual \( x \) projects on the adjectival scale from the initial point to the final one \( \text{fin} \). The initial point is the degree projected by the individual at the time of the event. The final point is a degree from comparison class \( P \) that falls on the other side of the \text{standout} standard of comparison. This requirement incorporates the intuition that \text{estar} predicates relate the present state of the individual to a potential one where it does not have that property. It is precisely the final point of the interval where the process comes into question in the case of open scale adjectives. Consider the examples in (134):

(134)  

a. Las hojas de los árboles están rojas.  

the leaves of the trees are\text{estar} red  

‘The leaves on the trees are red’

b. *Juan está famoso.  

juan is\text{estar} famous  

‘Juan is famous’

These are two of the prototypical examples that display the distribution of \text{estar} with open scale adjectives—leaves changing color has a natural process that drives the change, and therefore, the \text{estar} structure is valid, whereas being famous lacks that process, and therefore, the \text{estar}
predication crashes. It is proposed that this phenomenon is due to the fact that the comparison class for *famoso* ‘famoso’ does not provide the necessary degrees. The comparison class in *estar* predications has to be of the within-individual kind (Toledo & Sassoon, 2011; Gumiel-Molina et al. 2015). The formalization of the C morpheme that constructs the comparison class is, from example (124):

\[(135) \quad C = \lambda w(\_), \forall w'(\exists w'Aw)[(\text{x is realized as d at w'} \land P \text{ at d in w'})]
\]

The C morpheme is responsible, given a situation s, and all the possible accessible situations, for generating the appropriate degrees that are related to the present situation. The final point of the difference function $g^\uparrow$ in (135) has to be a member of that set of degrees in $P$. In the case of example (134a), the function $P$ includes such a degree—a degree where the leaves are not red—but not in the case of (134b)—the comparison class does not include any degrees where the individual is not famous—and therefore, the required interval cannot be generated. In the case where the open-scale adjectives and noun pairs have the required process, the accessibility relationship A can provide the necessary $w'$ situations from where the $P$ function can generate non-red degrees for the leaves. In cases that lack the process, the A relationship cannot access those situations, and therefore, those degrees are never included in the set defined by $P$. In other words, in (134a), we have access to situations where the leaves were not red, but in (134b), we cannot access situations where Juan was not famous.

This formalization presents an interesting question: why are those situations of a non-famous Juan inaccessible? It is trivially easy to think of examples of individuals that have acquired fame
and fortune during their lifetime. Why the opacity of those degrees? At this point, there is no clear answer as to why this should be the case. This phenomenon could be simply explained in terms of simply a lexical selection, which prevents adjectives such as rico ‘rich’ and famoso ‘famous’ from appearing in estar predicions. On the other hand, the characteristic distribution of these adjectives, all of them open-scaled ones cannot be ignored as a sign that there is a common thread that binds them together and accounts for their behavior regarding estar.

Furthermore, Chapter 4 in this dissertation takes the notion of the Natural Process presented in Section 3.6.2 of this chapter and, beyond its importance in the distribution of copulas, finds its importance on the online processing of said copulas in Spanish.

It is also worth mentioning in this respect that cases where real world knowledge does not have a reflection on the semantics are not without parallels in scalar accounts. For example, the adjective fast has an upper open scale, as the usual tests reveal:

(136) My car is {completely/absolutely/100%} fast.

Our knowledge of the world tells us that there is, in fact, the speed of light in a vacuum is the physical limit at which something can travel. Furthermore, this is not some obscure secret of quantum mechanics. It is common knowledge, as any grade school student can attest. Nevertheless, this knowledge has not transpired unto the semantics of the adjective, which remains an open-scale one, despite the evidence to the contrary. So, while the origin and characteristics of the processes that allow for certain open-scale adjectives to appear in estar
predications, but block others, certainly warrants further study, it should not be taken as a completely alien or unexpected feature of an scalar account of the copula distribution in Spanish.

3.7 Gradable Adjectives. Binary Scales

Thus far, the only adjectives that have been discussed are those that are traditionally called gradable. The concept of gradability (Sapir, 1944; Lyons, 1977), as it applies to adjectives, is the semantic property that enables a word to participate in comparative structures and to accept modifying expressions that act as intensifiers or diminishers of the quality expressed by the adjective. Those adjectives that admit comparatives and modifying expressions are said to be gradable, while those that cannot are non-gradable. In order to provide a comprehensive distribution of both Spanish copulas in adjectival contexts, it is necessary to include non-gradable adjectives. With this objective in mind, non-gradable adjectives are given a binary scale structure.

A binary scale is defined as a scale that is comprised as two single degrees. Each of these two degrees constitute the two truth conditions associated with the adjective. For example, given a non-gradable adjective such as atómico ‘atomic’, its scale could be represented as in (54):

(137) Atómico scale structure.
Under this consideration, *atómico* can now be considered a fully-closed scale adjective, with the caveat that only the minimal and maximal points of the scale available—the *atomic* and *non-atomic* degrees, and, therefore, it is non-gradable. The dashed line represents the empty scale between those two degrees. This approach to non-gradable adjectives does not allow for them to be incorporated into any situation that would require gradability, such as comparative structures. Yet, in the case of the present study, gradability does not play any role in copula distribution, only scale structure does. As such, non-gradable adjectives, if considered as a subset of closed scale adjectives, can be incorporated into the model. And crucially, when it comes to their copula distribution in Spanish, these binary adjectives behave exactly like their gradable, closed scale counterparts: if it can be conceivable that the subject of the predication would be in a different state than the one given, then a scalar interval can be determined and they can appear in *estar* structures. And since binary scale are also closed-scales, no natural process will ever be necessary.

In other words, binary adjectives that allow for the subject of the predication to access their other degree, will necessarily become *estar* predicates, while those that do not, will become *ser* ones. The example in (138) shows this distribution:

(138) a. Los dinosaurios {*son/están}* muertos.
   
   the dinosaurs {are*erp/are*estar}*dead

   ‘Dinosaur* are dead’

b. Los dinosaurios {eran/*estaban}* carnívoros.
the dinosaurs \{were_{ser}/were_{estar}\} carnivorous

‘Dinosaurs were carnivores’

Both *muerto* ‘dead’ and *carnívoro* ‘carnivorous’ are examples of non-gradable adjectives, but, in the case of dinosaurs being the subject of the predication, they have a critical difference—in the case of *muerto* ‘dead’, any living organism has potential access to the non-dead state, which is not the case for *carnívoro* ‘carnivore’. As predicted, *muerto* is a strict *estar* predication, while *carnívoro* is a strict *ser* one.

The inclusion of non-gradable adjectives unto the category of scalar adjectives provides a path to the account of the distribution of *ser* and *estar* with all adjectival complements. In the following sections, it will be shown that binary adjectives behave in exactly the same manner as their gradable, closed-scale siblings according to the present proposal.

### 3.8 Degree Phrases and *ser* and *estar* properties

The previous sections have detailed the interpretation of the predicates that will combine with either copula in Spanish and what requisites they must meet. The last step necessary is to formalize those requirements and incorporate them into a proposal for the Predicate Phrase that sits at the core of these predications. Based on the scalar structure of said predicates, the predicate phrase will take a Degree Phrase (Gumiel-Molina *et al.*, 2015) as its complement. At the DegP level, different properties of the *ser* and *estar* predications will be generated. The main
point of departure between Gumiel-Molina \textit{et al.} and the proposal presented in this dissertation lies in the nature of the properties that the DegPs that combine with \textit{ser} and \textit{estar} present. Gumiel-Molina \textit{et al.} propose that a \textit{estar} predications are based on comparing different degrees of the same individual, and \textit{ser} predications compare degrees of different individuals. In other words, the \textit{ser} and \textit{estar} distinction is determined by the comparison classes of the DegP. The current proposal, on the other hand, bases the \textit{ser} and \textit{estar} DegPs on the type of degree projected by the individual. While Gumiel-Molina \textit{et al.} can account for the data in \textit{ser} and non-evidential uses of \textit{estar}, their analysis falls short in accounting for some of the gaps found in the copular system in Spanish. They offer no proposal as to the lack of \textit{estar} \textit{rico} or \textit{estar} \textit{famoso}, or for the necessity of using \textit{estar} in closed-scale contexts such as \textit{el vaso está lleno} ‘the glass is \textit{estar} full’. Furthermore, it is unclear how certain evidential uses of \textit{estar}, such as \textit{el libro está genial} ‘the book is \textit{estar} great’, could be interpreted as comparing different stages of the same book.

With both \textit{ser} and \textit{estar}, a Degree Phrase is composed is around a covert morpheme (Kennedy, 2005, 2007; Husband, 2010) that specifies the requirements for a positive interpretation of the adjectival predicate. The main function of the \textit{pos} morpheme (von Stechow, 1984) is to relate the degree argument of the adjective to an appropriate standard of comparison (Cresswell 1977, Bierwisch 1989, and Kennedy 1999). It is assumed that the DegP present in a \textit{ser} predicate will encode a single degree on a scale, whereas the DegP in an \textit{estar} predicate needs to encode two points on the scale, and the \textit{pos} morpheme in both cases will need to be different. The following sections detail how these differences are formalized. As a reminder of the overall structure thus
far, the higher node proposed included the $V_{Cop}$ that contains the copula itself, as well as its complement PredP. Since the structure is the same both for *ser* and *estar* as in (139):

(139)

3.8.1 Degree Phrases in *ser* predication

In the case of Degree Phrases that can appear in *ser* predications, the *pos* morpheme is very similar to the one generally proposed in the literature (Kennedy 1997, Husband 2010) for English adjectival predication. As previously noted, this *pos* morpheme is of the form:

(140) $\lambda g(\epsilon, d) \lambda P(d, t) \lambda x(x). g(x) \succ standout$

For example, let us consider the following example involving *ser* in (141):

(141) Juan es inteligente.

Juan is *ser* intelligent
'Juan is intelligent’

In this case, the structure of the DegP will be as follows:

The pos morpheme, at the Deg node of the DegP, will first combine with the adjectival expression, which provides the appropriate dimension—intelligence in this case—and the scalar structure (Kennedy, 2005), which for ‘intelligent’ is an open one. Next, the appropriate comparison class needs to be integrated into the derivation. As previously mentioned, open scale adjectives require a *between-individuals* comparison class. The comparison class P is the set of degrees that the appropriate DP supplies. Since in our example there is no overt restriction, it is assumed that the individuals evaluated in the composition of the comparison class P will be a generic group (Bale, 2005; Toledo & Sassoon, 2011) encompassing all human beings—the desired interpretation of the example, lacking any further context is that he is intelligent for a human being in general. Given this comparison class, the function M will take it and generate the appropriate standout degree that human being generally must meet or
exceed in order to be considered intelligent. Finally, the DegP has the desired interpretation: given an individual, that individual will be considered intelligent provided his degree of intelligence is equal to or greater than the standard degree for intelligent human beings.

3.8.2 Degree Phrases in *estar* predication

In a similar manner to the one just described for *ser* predications, all of the necessary tools are not available to determine the derivation of an *estar* adjectival predication. At this point, the properties and structure of a DegP that will be able to appear in an *estar* predication can be completely stated. Let’s take a simple example, as in (142)

(142) Juan está alto

In this case, the adjectival expression *alto* ‘tall’ provides the function that maps individuals onto degrees on the tallness scale. The positive morpheme at the Deg node will require that function, which will also provide the necessary difference function, as well as the comparison class P,
which in this case takes the form of a within-individual comparison class, given that the derivation will need to have access to different degrees that the individual could project in different situations. The pos morpheme will first require that the individual project on the height scale that surpasses the standard of comparison determined by the function M, given the function alto’ and the comparison class P—in other words, the individual needs to be considered tall in this context. Furthermore, the difference function alto↑ will calculate the distance from the degree projected by the individual and a final point taken from the comparison class and will compare it to the distance from the degree projected by the individual to the standout degree of comparison and the former needs to be larger than the latter. This comparison between the two degrees generated by the difference function alto↑ is what ensures that we have access to such a degree that does not meet the standard of comparison. This is the precise quality that has been proposed as the defining feature of estar predications throughout the present study. At this point, the DegP can be fully derived, and can be applied as the argument of the PredP assumed for copular predications in Spanish.

3.9 Estar and the interpretation of open and closed scales

So far, my analysis of the DegP properties necessary for an estar predication has made no distinctions between open and closed scales. As previously pointed, this is a critical distinction in Spanish, since there is an asymmetry between these two classes of adjectives as it pertains to their copula distribution. The present implementation of the estar predications has no special provisions for closed vs open scale adjectives, although there are certain differences between
these two groups of adjectives worth mentioning, as they have consequences in the semantic
derivation of their interpretations.

Closed-scaled adjectives that can appear in an estar structure because of they have access to a
state with different truth conditions, must do so. For example, in (143):

(143)  a. El vaso {*es/está} lleno.
      the glass {isSER/isESTAR} full
      ‘The glass is full’

b. La puerta {*es/está} abierta.
   the door {isSER/isESTAR} open
   ‘The door is open’

These examples show cases of strict estar predications. It is not possible to represent the
situation of glasses being full or doors being open by means of ser. These examples, of course,
have been traditionally explained in terms of the predications representing transitory
characteristics of the subjects. Since one of the pervasive ideas exposed in the present study is
that neither ser nor estar depend on such considerations, and that they are just a consequence of
the underlying interpretation, the question as to why both examples in (143) have that
distribution must lie elsewhere.
Under a scalar account, the explanation to this exclusive distribution can be attributed to the principle of Interpretive Economy (Kennedy, 2007) stated in (144):

(144) **Interpretive Economy.** Maximize the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions.

This principle, as Kennedy (2007) interprets it, applies to the scalar structure of adjectives. According to Kennedy, the scalar endpoints in closed scale adjectives have a context-independent interpretation. For example, an empty glass will be interpreted as such in any conceivable context. Language users will resort to using these endpoints before they depend on contextually-dependent ones, such as the standard of comparison. Kennedy writes:

> although participants in a discourse may not be in full agreement about those properties of the context that play a role in the computation of context-dependent features of meaning, they are in agreement about the conventional meaning of the words and complex expressions in the sentences they use to communicate

(Kennedy 2007: 36).

Under this consideration, if a speaker is given a choice between producing a structure involving an adjectival predication that relies on endpoints and one that does not, the first one should be the preferred one, since endpoints represent conventional meanings that are not context-dependent. In the case of *ser* and *estar* with closed scale adjectives, there are such endpoints—supplied by
the comparison class P (Bale, 2010). As stipulated, the interpretation of the adjectival predicate centers around the positive morpheme at the core of the Degree Phrase. And, in the case of *ser* and *estar* predication, these *pos* morphemes have been defined as:

\[
\begin{align*}
&\lbrack \text{pos}_{\text{ser}}\rbrack = \lambda g \langle e,d \rangle \lambda P \langle d,t \rangle \lambda x \langle e \rangle . \ g(x) \Rightarrow M(g)(x) \\
&\lbrack \text{pos}_{\text{estar}}\rbrack = \lambda g \langle e,d \rangle \lambda P \langle d,t \rangle \lambda x \langle e \rangle . \ g(x) > M(g)(x) \land g^\uparrow_{\text{fin}} (x) \equiv g^\uparrow_{\text{standout}} (x)
\end{align*}
\]  

Given a closed scale adjective like *lleno* ‘full’ or *vacío* ‘empty’, the logical endpoints of the scale—the degree to which a certain container cannot be filled anymore, or the one when it cannot be emptied anymore—only enters the derivation as the *standout* degree of comparison. If an individual glass is to be considered strictly full, its degree of fullness must meet the standard of comparison, i.e. must be full. The *standout* degree is obtained from the comparison class P, which generates degrees from which the M function ranges over to determine that *standout* degree. However, the other degree that is necessary for the interpretation of the predicate is the g(x) degree that is projected by the individual at the time of the event. This degree can be on any point of the fullness scale, not necessarily the endpoint, and still yield a positive interpretation. In the particular case of an *estar* predication, the *pos* morpheme must necessarily use the endpoint of the scale when calculating the scalar difference. This is so because, in a closed scale, either the initial point or the final point of the interval projected by the individual must be that endpoint. The examples in (147) show this:
(147)  a.  El vaso está completamente lleno.
the glass is\textsc{estar} completely full
‘The glass is completely full’

b.  El vaso está medio lleno.
the glass is\textsc{estar} half full
‘The glass is half full’

c.  El vaso está un poco lleno.
the glass is\textsc{estar} a bit full
‘The glass is a bit full’

All of the predications in (147) yield a positive interpretation of the adjective \textit{lleno} ‘full’, although in a strict sense only in (147a) does the individual \textit{vaso} ‘glass’ projects the maximum degree on the scale. However, in all cases the lower endpoint of the scale—the degree that determines an empty glass—must enter the derivation as the final point of the derivation, as this is the only degree that determined a different truth condition, all the previous ones being full to some extent, and the endpoint being the only empty one. As stipulated, the final point of the derivation is a degree that the individual could potentially project and that has the opposite interpretation as the degree it currently projects. In all cases in (147), the initial point is a degree along the fullness scale that the glass currently projects, while the final point must be the degree where the glass is empty, i.e. the single degree that determines the lower end of the scale. We can see this graphically in (148):
In (148) the degrees of being a *little full* $d_3$, *half full* $d_2$ and *completely full* $d_1$ can all be compared with the endpoint of the scale, $d_4$, which represent the degree of being empty. All the intervals that $d_1$, $d_2$ and $d_3$ generate with $d_4$ are valid in order to generate an *estar* predication.

In this sense, *estar* predications of closed-scale adjectives require that the logical endpoints enter the interpretation of the individual regarding the properties of the adjective. Following the principle of Interpretive Economy, it follows that the *estar* option is preferred over the *ser* one, on these grounds. It is also worth pointing out that the relevant information to generate a *estar* predicate lies neither in the adjective nor in the subject of the predication, but in the comparison class. The comparison class will determine, by providing the necessary second point of the interval, whether the DegP will be able to appear in an *estar* structure or not.
As an interesting fact, this situation was not one that holds diachronically, as the following historical examples in (149) show (REAL ACADEMIA ESPAÑOLA: Banco de datos (CORDE) Corpus diacrónico del español. <http://www.rae.es> [2/25/2017]):

(149) a.  Señor, éste tu yantar es lleno de philosophía.

   lord, this your eating is\text{\textit{ser}} full of philosophy

   ‘Lord, your meal is full of philosophy’

b.  mostrasen algunas islas, de que es lleno aquel mar

   showed some islands, of which \text{\textit{ser}} full that sea.

   ‘they showed some islands, of which the sea is full’

Since the focus of the present study is synchronic in nature, the data in (66) must be seen as an interesting note, and an opportunity for further research. The only relevant fact presently is that closed scale adjectives predications that can establish an interval, and therefore generate an \textit{estar} structure, must do so in contemporary Spanish.

3.10 Evidential uses of \textit{estar}

3.10.1 Evidential and non-evidential uses of \textit{estar}

So far, the analysis has been centered on the traditionally called temporal uses of \textit{ser} and \textit{estar}. However, these are not the only cases that need to be accounted for. Already in Chapters 1 and 2, there was another context for \textit{estar} predication that was introduced: the evidential uses of \textit{estar}.
Any theoretical framework that strives to capture the operative factors in copula distribution in Spanish needs to necessarily take into consideration the evidential uses of *estar* (Mangialavori, 2013a; Camacho, 2015) section focuses on incorporating these cases into the theoretical framework presented in this chapter. The proposed analysis is that evidential uses of *estar* can also be subsumed under the general paradigm, and therefore are cases of predicates that need to refer to two different points on the scale that the adjective introduces. As such, evidential uses of *estar* also include a scalar interval in their interpretation.

Perhaps the most striking difference between evidential and non-evidential uses of *estar* in Spanish is that the former lack the temporal interpretation that has been suggested as the operative distinction in Spanish copula distinction. Camacho (2015) points out some of the features that these evidential uses display. One of the more salient differences between the evidential uses of *estar* and the ones described in previous sections is that the former do not seem to need to refer to other potential states of the individual subject of the predication. In fact, many examples of evidential *estar* do not allow for such interpretation, such as example (150):

(150) El nuevo disco de U2 está genial.

*the new record of U2 is* _estar_ *genius*

‘U2’s new album is great’

At first sight, sentences such as the one presented in (150) can be seen counterexamples to the account presented in this chapter so far. *Genial* ‘genius’ is an open-scale adjective, with no
discernible natural process behind it, and, moreover, it does not seem particularly susceptible to change. And yet, certain patterns can be identified that might shed some light on this particular use of *estar* and how it can be incorporated into the larger paradigm of copulas in Spanish.

The first striking difference between evidential and other uses of *estar* in Spanish is the latter being rather immutable predications. In example (150) it does not seem logical to think that a record would ever be able to change its quality, inasmuch as the music contained in it is permanent. This type of *estar* predication carries with it an interpretation of the concept of mutability of properties that is similar to the one previously observed in predicates that were barred from appearing in *estar* predications, as in (151), precisely because they do not allow access to other potentially different situations:

(151) Los tigres {son/*están} carnívoros.

\> the tigers {are\textsc{ser}/are\textsc{estar}} carnivorous

\> ‘Tigers are carnivorous’

One critical difference between (151) and our previous evidential *estar* (150), as Camacho (2015) points outs, is that the latter always allow for a *ser* variant, whereas the former are barred from *ser* predications. Naturally, this leads to the question what, if any, are the differences in interpretation between the sentences in example (152):
The proposed difference (Camacho, 2015; Mangialavori, 2013b) between (152a) and (152b) is that evidential uses of *estar* refer to a predication based more on personal experience and viewpoints than their *ser* counterparts (Escandell-Vidal and Leonetti, 2002)—they represent a higher degree of commitment on the part of the speaker towards the veracity of the predicate. Often, this translates to evidential uses of *estar* expressing a higher intensity than their *ser* counterparts. Going back to the examples in (152), it can be suggested that (152a) implies that the record is not simply good, but better than could be expected, whereas (152b) only indicates a general quality of goodness. This intuition is reflected in the number of evidential uses of *estar* that refer to the overall quality of the subject of predication.

This intuition is the basis that will allow for establishing a parallel, in scalar terms, between the non-evidential and evidential uses of *estar*. The following section addresses this task.
3.10.2 Evidential *estar* and scalar account

The starting point in the formalization of evidential uses of *estar* is that, just like in the case of non-evidential uses, the copula is compatible with certain predications and incompatible with others based on the presence or absence of certain properties in the PredP that is its complement, which includes a DegP that manifests such properties. As such, the main task is to determine what those properties might be and how are they manifested in the DegP. As has just been pointed out, evidential uses of *estar* carry with them a different interpretation of *estar* from the ones previously seen. The proposed formal analysis claims that evidential *estar*, much like their non-evidential counterparts, requires a predication that involves two degrees on the scale. The main difference lies in the nature of these two degrees. In the uses of *estar* seen in section 3.5 and 3.6, such as *Alberto está enfermo* ‘Alberto is sick’, these two degrees were each on opposite sides of the standard of comparison, as the requirement was that the individual could potentially project a degree that did not meet the standard of comparison. In the case of evidential *estar*, the two required degrees are determined, on the one hand by the degree projected by the individual at the time of the event, and on the other hand, just as in the non-evidential uses, by the standout standard of comparison. As previously mentioned in the discussion of (152), an evidential *estar* predicate requires that the degree projected by the individual is not just one that exceeds the standard of comparison, but also does so in more than a determined amount, which indicates a higher degree of speaker commitment to the proposition. This leads to the conclusion that there is a different positive morpheme in the DegP contained in the evidential uses of *estar*. Formally, the positive interpretation morpheme that controls this type of predication can be expressed in the terms seen in (153):
The pos morpheme takes a degree function \( g \), provided by the adjectival expression, a comparison class \( P \), comprised of a set of degrees, and an individual \( x \), and establishes that the degree projected by the individual on the scale determined by the adjective is greater than the standout degree of comparison, as determined by the function \( M \), given the adjectival function and the comparison class. Furthermore, the difference function \( g \uparrow \) will determine the distance between the standout degree and the degree projected by \( x \), which needs to be greater than a standout distance, as generated by the function \( N \). This function \( N \) is similar to the function \( M \), in that it takes an adjective function \( g \), a comparison class \( P \) and an individual \( x \), only instead of returning a standout in the form of a single degree, it returns a standout interval of comparison for the present situation.

In essence, the degree that the individual projects must not only be above the standard of comparison, but must surpass it by a certain quantity. Graphically, this could be represented as follows:

\[
\text{\#pos} = \lambda g_{(c,d)}\lambda P_{(d,t)}\lambda x_{(e)} . \ g(x) \geq M(g)(P)(x) \land g\uparrow_{\text{standout}}(x) \geq N(g)(P)(x)
\]
As it is indicated in the graph, on the scale $S$, the distance from the degree of the individual, $g(x)$, to the standout standard of comparison, is greater than $N(g)(P)(x)$, the distance required for an evidential use of *estar* in this context. Applied to the example in (152a), $g(x)$ is the degree of genius that U2’s latest record possesses. The standout represents the degree that degree that a record must exceed in order to be considered genius. The $g_{standout}(x)$ degree expresses the interval of distance between how genius U2’s record is and how much any record needs in order to be considered genius. And $N(g)(P)(x)$ represents the distance to the standout that needs to be exceeded in order for a record to be able to appear in an evidential *estar* structure. The interpretation of (69a), in terms of scalar structure, could then be translated as *I consider U2’s new record is genius enough that it surpasses the minimum requirement of a genius record by a determined amount*. In this manner, both evidential and non-evidential uses of *estar* can both be give a common

3.10.3 Evidential *estar* and comparison classes

Another difference between evidential and non-evidential uses of *estar* in adjectival predication lies in the nature of the comparison classes that their predications require. In the case of non-evidential *estar*, the comparison class necessarily had to include within-individual degrees of comparison—the connection to a degree on the other side of the standard of comparison requirement. In the case of evidential uses of *estar*, this access to other degrees that the individual subject of the predication could project is not a requirement. As such, the comparison class can be strictly between-individuals with no repercussion to the positive interpretation of the
predication. This effect can be seen in cases where evidential *estar* structures are of a permanent nature, with no possibility of change:

(154) Esta película está entretenida.

this movie is<sub>ESTAR</sub> entertaining

‘This movie is entertaining’

It is hard to imagine a situation in which a particular movie would mutate from an entertaining state to a non-entertaining one. In this manner, the interpretation of (71) is close to being of immutable properties, and incapable of appearing in *estar* predications as some previously seen examples—dinosaurs being carnivorous, for instance. It then follows that in (71), the comparison class cannot possibly contain different instances of the movie having different degrees of entertainment qualities, and thus the comparison class is of a between-individual nature, comprised of different movies with different degrees along the scale determined by the adjective.

With the positive morpheme and the comparison class defined, it is now possible to establish the derivation of an example of evidential use of *estar*, for example in sentences like (155):

(155) a. El nuevo disco de U2 está bueno.

the new record of U2 is<sub>ESTAR</sub> good

‘The new record of U2 is<sub>ESTAR</sub> good’
In this case, the adjective *bueno* ‘good’ provides the measuring function that will project individuals as degrees on the scale of goodness. The *pos* morpheme at the Deg node will take the adjective as its complement, and, together with the comparison class P, will generate the standout degree of comparison, which the degree projected by the individual $x$ needs to exceed. Furthermore, the distance on the scale, as calculated by the difference function $g^\uparrow$, from the degree of the individual and the standout needs to be bigger than a standard difference, calculated by the function $N$ for the given comparison class $P$ and the individual $x$. This produces an interpretation in which the individual $x$ will not only need to be considered good in this context, but also will need to exceed the standard in a certain quantity. And, as it is the case in non-evidential uses of *estar*, the PredP is generated with this DegP as its complement, which in turn will be the complement of the $V_{Cop}$P. As in non-evidential uses of *estar*, the appropriate *pos* morpheme generates the structure with the required interpretation—an interval along a scale. The main difference between the evidential and non-evidential uses of *estar* lies then not in the presence of this interval, but in the manner it is constructed. In the case of non-evidential *estar*,
by accessing a degree with a different truth-condition, and in evidential uses, by distance to the standard of comparison.

The final interpretation of the predication can be given as *El nuevo disco de U2 está bueno* at some time iff it *is the case that U2’s new record holds the property, at a point in time prior to the time of the event, such that the degree of goodness that the record projects is to equal or greater than the standard of goodness for that particular situation and surpasses it in a certain amount.*

3.10.4 Evidential *estar* and Interpretative Economy.

In the case of evidential uses of *estar*, as Camacho (2015) points out, the predicates are always capable of appearing in *ser* predications as well. The scalar account just presented predicts and explains this behavior. Given that the comparison class in evidential uses of *estar* needs to be of a between-individuals nature, only relative adjectives, such as bueno ‘good’ or grande ‘big’ (Kennedy, 2005; Bale, 2010), will be able to appear in this type of evidential *estar* structure. As previously determined, an absolute adjective, such as vacío ‘empty’ or cerrado ‘closed’ will have a standout degree of comparison that is an endpoint on the scale, providing a contextual free interpretation. As such, these endpoints provide a cognitively salient (Potts, 2008) degree on which the interpretation will latch on first, as per Interpretive Economy (Kennedy, 2007) will require an interval interpretation over a single point one. In the case of evidential *estar*, the interpretation also hinges on the standard of comparison in order to generate the interval, but the evidential nature of the interpretation allows for an optional nature of the particular use. Hence,
Camacho’s observation about the systematic availability of *ser* and evidential *estar* predications in the same structures—although with slightly different meaning—should not be surprising.

### 3.11 Conclusions

In this chapter, the intuitions about the interpretation of both *ser* and *estar* have been implemented in a formal framework that can account for their distribution in adjectival contexts as seen in Chapter 1. Based on scalar considerations, the temporal restriction that has plagued classical accounts has been avoided, and, instead a broader, more accurate requirement (Romeu, 2105), based on potential degrees has been implemented. Furthermore, a hitherto unobserved phenomenon has been accounted for and integrated in the proposed system—open scale and closed scale adjectives impose different conditions regarding their copula distribution, and, therefore, scale boundaries need to play a role in any formal, semantic account of *ser* and *estar* in Spanish. It is also worth pointing out again that it is not the scalar structure of the adjective that determines whether a predicate can combine with *estar*, but rather the information contained in the comparison class. For non-evidential uses of *estar*, this comparison class needs to include a degree with a different truth-condition from the one expressed in the predication. For evidential uses of *estar*, the comparison class needs to provide the scalar distance that the degree projected by the individual must surpass. In essence, it is not the lexical information contained in the adjective that determined whether it can appear in *ser* or *estar* predications, but rather the pragmatic associated with that adjective and that individual. Scalar structure on its own cannot be used as a predictor of copula choice. This can be seen in examples where changing the subject of the predication will change the copula, such as an adjective like *cerrado* ‘closed’ appears with
 estar in the case of describing a window, but with ser in the case of describing the set of prepositions in Spanish.

Although scalar structure per se cannot determine copula choice, a framework that incorporates scalar structure considerations such as the one in the present study can therefore account for the copular data that is present in the language, but also account for the apparent gaps in the system—why adjectives rojo ‘red’ can appear in estar predications, but adjectives like rico ‘rich’, famoso ‘famous’ or necesario ‘necessary’ cannot.

Furthermore, the evidential uses of estar have also been accounted for by the present approach. Despite their common use in the language, they have traditionally been ignored in the literature. It has been shown, however, that regardless of their difference in interpretation with non-evidential uses of estar, they can be accounted for under the general paradigm of estar predications requiring a scalar interval. Although different provisions need to be made regarding the manner in which these evidential uses generate their scalar interval, the fact that they can be included under the same umbrella as their counterparts gives the whole system a certain cohesiveness.

Of particular relevance is the importance of the process driving the presence of estar in open scale adjectival predication. As has been pointed out, noun-adjective pairs that lack this natural process cannot generate the appropriate comparison classes, and therefore, cannot produce the necessary scalar interval in order to appear in estar structures. Although at this point there is no simple way in which to define the nature of this process, or how it comes to be found in some
adjectival contexts but not others, further investigations can be made about the impact that these processes have on the copular system in Spanish. Chapter 4 follows up on this idea by trying to determine what, if any, are the effects of these processes on the online interpretation of the copula *estar* in Spanish. As argued in this chapter, copulas in Spanish project a requirement on their predicates—a single degree for *ser*, and an interval for *estar*. The experiment presented in Chapter 4 investigates the incremental processing of copular sentences, and specifically, whether these requirements projected by the copulas can manifest as anticipatory fixations in an eye-tracking experiment.
CHAPTER 4. Online Processing of *Ser* and *Estar*

Section 4.1 Overview

In Chapter 3, I constructed a theoretical framework for the distribution of *ser* and *estar* based on the semantic properties of the predicates they combine with. One of the ideas put forward in that chapter is that *estar* is sensitive to the scalar structure of its predicates. More specifically, open-scale adjectives that combine with *estar* not only need to determine an interval on the adjectival scale, which is a requirement for all adjectives regardless of scale structure, but also need to have a natural process responsible for the change in the quality defined by the adjective.

The natural process is critical for generating *estar* structures with open-scale adjectives. As mentioned in Chapter 3, once the process is assumed to be absent from the predication, *estar* is no longer acceptable. See example (156):

(156) a. Alberto tiene diez años y está alto.
   Alberto has ten years and is tall
   ‘Alberto is ten years old and is tall.’

b. #Alberto tiene cincuenta años y está alto.
   Alberto has fifty years and is tall
   ‘Alberto is fifty years old and is tall.’
In (156a), the topic of the predication is a child, who can be assumed to be going through the process of growth. Therefore, the combination of *estar* with the open-scale adjective *alto* ‘tall’ is acceptable for the interpretation that he is tall for his age. In (156b), the natural process of growing is assumed to have ceased for a 50-year-old man, and the combination of *estar* with *alto* ‘tall’ results in an odd sentence. It is precisely the relevance of such a process that this chapter focuses on.

While Chapter 3 argued for the importance of the natural process in *estar* predication with open-scale adjectives from a theoretical perspective, the present chapter aims to provide empirical support for the predications that stem from this approach. In particular, this chapter presents an eye-tracking experiment that examines the timing of the interpretation of copulas during online processing of spoken sentences that contain such structures. Section 4.2 reviews two previous experimental studies, Schmitt & Miller (2007) and Holtheuer et al. (2012), about *ser* and *estar* with open-scale adjectives. Section 4.3 motivates the use of the visual world paradigm and discusses how the experimental outcome contributes to the approach presented in Chapter 3. Section 4.4 describes the experimental design and the materials, and Section 4.5 presents the results of the experiment. Section 4.6 discusses the implications of the experimental results.

### 4.2 Past experimental data

The main objective of this chapter is to provide empirical data of how the presence or absence of a natural process affects the interpretation of the copula *estar* in Spanish. The experiment presented in this chapter is the first to reveal the way in which Spanish copulas are processed.
during spoken sentence comprehension. Although the requirement of a natural process for a predication with *estar* and open-scale adjectives has never been investigated with an experimental method, some studies have provided experimental data on the *ser* and *estar* distributions in Spanish. Two of those studies have focused on open-scale adjectives.

### 4.2.1 Schmitt & Miller (2007)

Schmitt and Miller (2007) investigated children’s ability to interpret the temporal restrictions that *ser* and *estar* impose on predicates, Stage Level and Individual Level predicates, discussed in Chapter 2. Schmitt and Miller build on the idea that the distinction between *ser* and *estar* is ultimately an aspectual distinction and aimed to determine whether children can process the permanent vs. temporary implicatures associated with *ser* and *estar*. Schmitt and Miller predicted that while *estar* predicates are always interpreted as states of an individual, *ser* is flexible in terms of its event type properties, and can, depending on the context, be interpreted as a temporary state or a permanent one. Accordingly, both *ser* and *estar* can be interchangeably used by children to describe the same situation. Second, *estar* predicates have an implicature of temporariness associated with them, both for children and adults. Third, when *estar* is combined with an adjective, its interpretation requires the identification of the relevant period of time in which the property described by the adjective holds. The experiments tested whether children show preferences between the two copulas in contexts where both are logically possible.

The first experiment in Schmitt & Miller (2007) tested children’s comprehension of the copulas *ser* and *estar* through an elicitation task. Children were shown pictures of people with body parts
that exhibited permanent and temporary colors (e.g. red tongue versus green tongue), and were asked to complete sentences about the characters in the pictures. The example in (157) provides one of the sentence sets the children were asked to complete.

(157) Este es Manolo. Su lengua se puso verde por que estaba tomando este jugo de kiwi, pero el tiene la lengua roja.
This is Manolo. His tongue turned green because he was drinking this kiwi juice but he has a red tongue (picture of a red tongue with green shading on the top of it).

a. La lengua de Manolo es _______.
   “Manolo’s tongue is .”

b. La lengua de Manolo está _______.
   “Manolo’s tongue is .”

The experiment was conducted with 35 monolingual children, ages 4;5-6;3 (mean of 5;3), and with 24 monolingual adults. With the sentence completion task, adults chose the permanent color (e.g. red) after ser 94% of the time and the temporary color (e.g. green) for estar 100% of the time. In the case of children, those numbers were 82% for ser and 85% for estar. The discrepancies between adults and children’s scores were significant for the estar condition, but not for the ser condition. These results suggest a difference in the way children 4;5 of age interpret estar, and that children “are calculating the implicatures” associated with the distinction.

The second experiment of Schmitt & Miller (2007) aimed at determining whether children can use discourse information or whether they rely on world knowledge in their interpretations of ser
and *estar*. In this experiment, 24 Spanish-speaking children, ages 4;7–6;0 (mean age: 5;5) and 20 Spanish-speaking adults were given a picture matching task. They were given short stories about animals who had eaten magic beans and had subsequently changed size for a short period of time (See Fig. 1).

![Sample target pictures and story from Schmitt & Miller (2007)](image)

After reading the stories and seeing the accompanying pictures, participants answered two questions, one with *ser* and one with *estar*, regarding the animals in the story. Example (158) provides one such set of questions:

(158) a. ¿Cuál gato está gordo? ESTAR

which cat is *ESTAR* fat

‘Which cat is fat?’

b. ¿Cuál gato es gordo? SER
which cat is $\text{ser}$ fat

‘Which cat is fat?’

Given the picture and context in Fig. 1, the expected adult response for the $\text{estar}$ condition (158a) was the cat on the right—i.e. the cat that had grown fat after eating the beans—and for the $\text{ser}$ condition (158b), the cat on the left—i.e. the cat that was originally fat. For the question involving $\text{ser}$, adults were expected to prefer the cat on the left. In the $\text{ser}$ condition, adults chose the permanently fat cat 95% of the time. However, in the $\text{estar}$ condition, they only chose the temporary cat 50% of the time. The pattern was different for children. In the $\text{ser}$ condition, children chose the permanently fat cat 42% of the time, the temporarily fat cat 35% of the time and both cats 23% of the time. In the $\text{estar}$ condition, children chose the temporarily fat cat 77% of the time and the permanent 23% of the time. These results are summarized in Figure 2.

![Figure 2. Results Experiment 2 from Schmitt & Miller (2007). Percentage of copula used per condition.](image-url)
While adults were fairly consistent in their interpretation of *ser* and chose the cat with the permanent condition, they were not do so in the *estar* condition. In the case of children, the response pattern is reversed; they dominantly chose the cat in a temporary status for *estar*, but they did not show a strong preference for the *ser* condition.

Based on these results, Schmitt & Miller reached two conclusions: first, when children can rely on their experience and do not need to choose a time interval, their preferences are adult-like 80% of the time. As the context becomes more complex, children seem to demonstrate the knowledge that, while *ser* can be used in any context, *estar* demands a temporal interpretation.

These results, however, can be interpreted in a different way if considered under the framework proposed in Chapter 3. The framework proposed in this dissertation establishes that *estar* predications with open-scale adjectives require that the property described by the adjective can change for the subject of the predication, and that potential change is driven by a natural process. All of the adjectives that Schmitt & Miller (2007) employ in their experiments are open-scaled adjectives—size, height, color—and therefore will require an active, valid process in order to appear in an *estar* predication. It should not be surprising then that adults would perform better in the *ser* condition, because they do not need to consider the presence or absence of a natural process, and just need to concentrate on the ability of the individual to change the property in question. In the *estar* condition, on the other hand, they are presented with a highly unlikely or conditional process—eating magic beans—which may have counted to some participants as a valid process, but not to others. Children may not have learned the importance of the natural
process for verifying predicates with *estar*, and thus any interpretation of the data in Schmitt & Miller (2007) must be taken with a grain of salt. In any event, one thing seems clear: the distribution of *estar* in contexts like the ones provided in these two experiments is not as straightforward as it might seem, and it certainly poses, even to adult native speakers of Spanish, a challenge.

### 4.2.3 Holtheuer et al. (2012)

Holtheuer *et al.* (2012) is an extension of Schmitt & Miller (2007) and it is also aimed at determining the different interpretations of *ser* and *estar* in children and adults in open-scale adjectival contexts. Holtheuer *et al.* base their interpretation of *ser* and *estar* on different temporal properties: *estar* being a predication of a period of time and *ser* having no such restrictions. Two hypotheses were tested in 2 experiments: *ser* and *estar* are not in complementary distribution and certain situations may be described by both copulas. *Ser* is the more transparent of the two copulas, having no temporal restrictions. *Estar* imposes stricter requirements on the predication, as it carries a presupposition that the predication is restricted to a temporal slice in the life of the individual. Since *ser* is given a smaller number of restrictions, if the predication does not involve any overt temporal restrictors, *ser* should be used more than *estar*.

In their first experiment, Holtheuer *et al.* examined children’s production of *ser* and *estar* with open-scale adjectives in contexts that portrayed either temporary or permanent characteristics of individuals. Twenty-six Spanish-speaking children ages 3;10 to 7;4 (mean age 5;7) and 25 native
Spanish-speaking adults participated in an elicited production task. Participants were presented with a short story, similar to the ones used in Schmitt & Miller (2007) and seen in section 4.2.2, about a character or a set of characters that had specific properties related to size (e.g., they were born fat, or they were born short) that changed as the story progressed. Then, participants were asked to describe both the original and new properties. The distribution of *ser* and *estar* in their answers was then examined. In each story, either the same character underwent the change, or a different character did. See Figure 3 for an example:

![Fig. 3. Same character and different character condition. From Holtheuer et al (2011)](image)

The authors predicted that the *ser* copula would be favored in order to describe the same characters in the first condition, and *estar* in the case of the second condition. In the case of describing two different cats, *ser* is predicted in both cases. The results, both for children and adults are summarized in Table 2.
Table 2. Experiment 1 results from Holtheuer et al. (2011). In the Ø option, no answer was given.

The results indicate that children are adult-like in the different character condition: they are biased to produce *ser* almost 100% of the time. In the same character condition, however, children and adults did not behave in the same manner. Adults produced *ser* in contexts both of permanent and transient properties much more than children. In other words, the same character condition did not strongly bias adults for *estar*. In contrast, children showed a strong preference for *estar* in this condition. As in Schmitt & Miller (2007), children and adults exhibit different results in their use of *estar* in the presence of open-scale adjectives depicting transient properties of individuals.

The second experiment in Holtheuer *et al.* (2007) used color adjectives, instead of physical size. Additionally, this second experiment had a between-subjects design, comparing 2 groups of children and 2 groups of adults. Twenty-seven Spanish-speaking children (4;5 to 7;4, mean age 5;6) and thirty-four Spanish-speaking adults participated. Fifteen children (4;8 to 7;4, mean age 5;5) and thirteen adults were tested on the different character condition and twelve children (4;5
to 7;0, mean age 5;8) and 15 on the same character condition. The results are summarized in Table 3:

<table>
<thead>
<tr>
<th></th>
<th>Different Character (DC) Condition</th>
<th>Same Character (SC) Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adults</td>
<td>Children</td>
</tr>
<tr>
<td>SER + SER</td>
<td>99 (42/43)</td>
<td>86 (43/50)</td>
</tr>
<tr>
<td>SER + Ø</td>
<td>1 (1/43)</td>
<td>0 (0/50)</td>
</tr>
<tr>
<td>TOTAL SER</td>
<td>100 (43/43)</td>
<td>100 (50/50)</td>
</tr>
<tr>
<td>ESTAR + ESTAR</td>
<td>0 (0/43)</td>
<td>0 (0/50)</td>
</tr>
<tr>
<td>ESTAR + Ø</td>
<td>0 (0/43)</td>
<td>0 (0/50)</td>
</tr>
<tr>
<td>Total ESTAR</td>
<td>0 (0/43)</td>
<td>0 (0/50)</td>
</tr>
<tr>
<td>MIXED COPULAS</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3. Experiment 2 results from Holtheuer et al. (2011)

The results are very similar to those of the first experiment. Both children and adults favored *ser* in the different character condition, but the production of both copulas in the same character condition again exhibited the reversed distributions between children and adults. Holtheuer et al. therefore concluded that neither children nor adults use *ser* and *estar* in a complementary or mutually-exclusive manner. The results of this experiment may be explained with the effect that the natural process has on the distribution of *ser* and *estar* with open-scale adjectives. In order for *estar* to be used for a color change within the same character, the change due to the magic beans, must be considered as a reasonable natural process. The differences between adults and children may reflect the difference in the comprehension of the world. While asults may have rejected the use of *estar* for the change due to the magic beans, children, who may not have acquired this requirement, may have focused more on the temporality of the event.
Schmitt & Miller (2007), and Holtheuer et al. (2011) both tested the effect of the property of the predication on the choice of Spanish copula. The asymmetry remains between the use of *ser* and *estar*. The context for use of *ser* with an open-scale adjective seems clearer than the context for the use of *estar*: the predication requires an inherent, stable property. Both children and adults favor *ser* in cases where the property defined by the adjective was one that the subject of the predication would not be able to change. This maps well with the general proposal of this dissertation that *ser* predicates indicate a single point on a scale with no access to a different state. In temporary readings, the choice of *estar* may depend on the participants’ interpretation of the type of process that drives the change. In both studies, children may have focused on the effect of the magic beans, which may have highlighted the temporariness of the change in size. They also may have associated the overt visual change of the same character with the effect of the magic beans and thus exhibited the robust preference of *estar* for such a case. Adults may have focused more on the consequence of eating the beans and may have entertained the possibilities that the size after eating the beans can be a temporary property whether or not the cat had been inherently fat (Schmitt & Miller, 2007) and that the cat that changes size due to the magic beans may remain fat (Holtheuer et al., 2011). In either case, the scenario that involves fantasy can lead to a variability in the interpretation behind the predication, making it difficult to attribute the group differences in the data either to the difference in the lexical knowledge or to the difference in the experiment-specific task interpretation.

Although the acquisition of Spanish copula needs to be further explored, it would require a more solid set of empirical data from adults that establishes a reference for the developmental
trajectory. The empirical research in this dissertation therefore examines the distribution of copula in Spanish in adults, and tests whether adult Spanish speakers rely on the copula in predicting the content of upcoming predication.

4.3 Online Processing of *ser* and *estar*

The core research question of the present experimental study is whether *estar* evokes the interpretation of a natural process during online comprehension of spoken Spanish sentences. The information obtained from online processing of *ser* and *estar* predications should allow us to determine what the factors that affect the interpretation of the copulas are while the sentence unfolds. Critically, online responses would reveal whether the subject noun phrase and the following copula are sufficient to project the interpretive constraints (such as natural process) or listeners must hear the rest of the predication to evaluate the proposition. Offline data, such as those from Schmitt & Miller (2007) and Holtheuer *et al.* (2011), are unable to provide this type of introspection.

Previous research with an online eye-tracking technique has shown that listeners build expectations about the upcoming content of the sentences according to the syntactic and semantic information of the words that have unfolded (Altmann & Kamide, 1999; Kamide *et al.*, 2003). The present experiment tests if such incremental process is triggered only by the subject and copula in spoken Spanish.
4.3.1 Visual World Paradigm

Past studies on incremental processing demonstrated how lexical information that unfolds in an early part of a sentence leads to a prediction for the rest of the sentence (e.g., Altmann & Steedman, 1988; Altmann, Garnham, & Henstra, 1994; Mitchell, Corley, & Garnham, 1992; Tyler & Marslen-Wilson, 1977; Altmann & Kamide, 1999, 2007; Kamide et al., 2003). Many of these studies have made use of an eye-tracking technique called visual world paradigm (Cooper, 1974; Tanenhaus, Spivey, Eberhard, and Sedivy, 1995), in which the gazes to particular visual referents at any moment are linked to the anticipation that the listener has built based on the information gathered so far. Tanenhaus and Trueswell (1995) highlight the methodological advantages of this paradigm regarding spontaneity of eye-movements, naturalistic task environment and the unconscious immediate responses to linguistic input that cannot be captured by traditional offline tasks.

In these visual world experiments, participants are typically presented with a visual scene or a set of visual objects, and an auditory input that may be related to objects in the scene. Participants’ eye movements across the visual scene are recorded while the auditory input unfolds. The use of realistic scenes allows researchers to assess how the listeners’ perception of the scene and their world knowledge about scenes or events affects their understanding of the spoken utterances (Henderson & Ferreira, 2004). An advantage of the visual world paradigm compared to other psycholinguistic paradigms is that the listeners do not have to perform any meta-linguistic judgments. The visual world paradigm solely relies on the listeners' tendency to look at relevant parts of the display as soon as any information associated with the visual target becomes
available (Huettig et al., 2011). Because of its fine-grained temporal resolution, the paradigm has been applied to a wide range of research topics, such as spoken word recognition (Allopenna et al. 1998; Dahan et al., 2001; Huettig & McQueen, 2007; McMurray et al., 2008; Wiener & Ito, 2015), online computation of verb-thematic information (Altmann & Kamide, 1999; 2007, 2009), pronoun resolution in discourse (Arnold et al., 2000), effect of visual context on referential resolution (Chambers et al., 2003; Trueswell et al., 1999; Spivey et al., 2002) and effect of prosodic prominence on referential resolution (Dahan et al., 2002; Ito & Speer, 2008; Ito et al. 2012; 2014; 2017).

Dahan and Tanenhaus (2004) examined the processing of lexical competitors in sentential contexts using naturally occurring Dutch constructions in which the verb precedes its subject noun phrase. The subject noun phrase followed either a main verb that placed strong semantic constraints on its subject (e.g., Never before climbed a goat so high) or an auxiliary or modal verb that did not place such constraints on its subject (e.g., Never before has a goat climbed so high). The authors showed that when the preceding verb context established thematic constraints, eye fixations were limited to the visual referent that matched the constraints. Listeners accessed lexical meanings and integrated them with the relevant context from the earliest moments of lexical processing.

Altmann and Kamide (2003) showed that the incremental interpretation of a sentence is based on the partial interpretation of what has been encountered thus far. They examined how the semantic interpretation of a verb can cause the anticipation of an upcoming Theme. For example,
the full interpretation of a sentence fragment such as *The boy will eat* must evoke the anticipation that something will be eaten. Their experiments showed how in a sentence such as *The woman will spread the butter on the bread*, participants showed anticipatory looks towards the appropriate Goal picture, *the bread*, while the Theme, *the butter*, was unfolding. This shows how the subject anticipates, either at the verb or during the first post-verbal argument (the Theme), information pertaining to the second post-verbal argument (the Goal).

Both these experiments, Dahan and Tanenhaus (2004) and Altmann and Kamide (2003) and many other studies with the visual world paradigm have demonstrated incremental interpretation in anticipatory eye movements during online spoken sentence comprehension.

### 4.3.2 Why Color Adjectives

The visual world paradigm employed in the present study focuses on color expressions as the target predications with open-scale adjectives. This decision is motivated by the previous observation that the color contrast is more swiftly grasped than the size contrast (Ito & Speer, 2011), and the ease of collecting visual stimuli materials that involve natural processes. The critical sentences that the subjects heard in the experiment included either the *ser* or *estar* copula, followed by a color adjective. Color adjectives provide an ideal context for the study of copulas in Spanish, as they may combine both with *ser* and *estar*. See example (159)

(159) a. La bicicleta *{es/#está}* roja.

- the bicycle *{isSER/isESTAR}* red

‘The bicycle is red.’
b. Las hojas {son/están} rojas.
the leaves {is/estar} red

‘The leaves are red.’

In (159a), *estar* results in an unacceptable sentence based on the lack of a natural process driving the change in color. In (159b), the predication may refer to a natural process and *estar* is a valid copula choice.

In order to confirm that Spanish does indeed favor the presence of a natural process in *estar* with color adjectives, a corpus search was carried out. This search made use of the CREA and Davis corpora of Spanish, and the search criteria included six adjectives: *rojo* ‘red’, *amarillo* ‘yellow’, *azul* ‘blue, *verde* ‘green’, *blanco* ‘white’ and *negro* ‘black’. The search included these adjectives in all their possible gender and number forms and all possible person and number combinations of *estar* in Present Indicative, Present Subjunctive, Preterit, Imperfect and Future tenses. The adjective expression could also appear within the five words preceding or following the copula. Idiomatic expressions such as *estar negro* ‘to be black’ meaning being obfuscated or *estar verde* ‘to be green’ meaning being a novice or not experienced at something were excluded from the analysis. Table 3 summarizes the findings of this corpus search.

<table>
<thead>
<tr>
<th>Context</th>
<th>Number of cases (874 total)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>437</td>
<td></td>
</tr>
<tr>
<td><em>Estar + Natural Process</em></td>
<td>422</td>
<td>96.5%</td>
</tr>
<tr>
<td><em>Estar + Artificial</em></td>
<td>15</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

Table 4. Distribution of *estar + color adjectives across the CREA and Davis corpora.
The results of the corpus search confirm the intuition put forward in the theory: Spanish speakers heavily favor the use of *estar* color expressions with natural processes over artificial processes. The examples in (160) are taken from the CREA corpus of Spanish:

(160) a. Al salir a la calle había anochecido y el cielo *estaba rojo* como la bóveda del infierno.

‘When we got out on the street, the sun had come down and the sky was red like the vault of hell.’

b. Viene de la mina. […] La mayor parte de sus manos y cara está negra.

‘He comes from the mine […] Most of his hands and face are black.’

4.4 The Experiment

4.4.1 Participants

Subjects for the experiment were recruited primarily from the Ohio State University student and employee community. They were recruited via personal contact and flyers in public boards. Fifty one participants were recruited, all native speakers of Spanish over the age of 18 who have completed at least their secondary education in Spanish. In order to ensure their level of Spanish, a small, informal conversation in Spanish was carried out with each participant before scheduling
an experiment. In this interview, they commented on aspects of their everyday life and were also asked about the color of certain very day items.

All participants answered a linguistic background questionnaire inquiring about their education level, number of hours they use Spanish per day, dialectal background and language background of parents. Table 5 summarizes their responses on this questionnaire:

<table>
<thead>
<tr>
<th>Category</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>19 to 41. Avg: (29.04)</td>
</tr>
<tr>
<td>Hours Spanish per day</td>
<td>5.32 Avg.</td>
</tr>
<tr>
<td>Months in the US</td>
<td>17.54 Avg.</td>
</tr>
<tr>
<td>Dialect of Spanish</td>
<td>Spain (14)</td>
</tr>
<tr>
<td></td>
<td>Mexico (14)</td>
</tr>
<tr>
<td></td>
<td>Bolivia (4)</td>
</tr>
<tr>
<td></td>
<td>Colombia (3)</td>
</tr>
<tr>
<td></td>
<td>Venezuela (2)</td>
</tr>
<tr>
<td></td>
<td>Puerto Rico (5)</td>
</tr>
<tr>
<td></td>
<td>Peru (2)</td>
</tr>
<tr>
<td>Highest Degree of Education Achieved</td>
<td>Secondary (14)</td>
</tr>
<tr>
<td></td>
<td>Undergraduate (30)</td>
</tr>
<tr>
<td></td>
<td>Master (6)</td>
</tr>
</tbody>
</table>

Table 5. Background information of all participants.
Data from participants who did not perform above 50% in the picture selection task in the eye-tracking experiment were excluded from further analysis. About half of the recruited participants were screened out by this criterion, suggesting that the task was not particularly easy even for the native Spanish speakers (see below for further discussion). Data from twenty-eight out of the fifty-one participants whose accuracy rate was above 50% are reported below. The background data for those twenty-eight participants is summarized in Table 6:

<table>
<thead>
<tr>
<th>Category</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>21 to 41. Avg: (29.3)</td>
</tr>
<tr>
<td>Hours Spanish per day</td>
<td>5.44 Avg.</td>
</tr>
<tr>
<td>Months in the US</td>
<td>21.54 Avg.</td>
</tr>
</tbody>
</table>
| Dialect of Spanish              | Spain (12)  
                              | Mexico (11)  
                              | Bolivia (2)  
                              | Colombia (1)  
                              | Venezuela (1)  
                              | Puerto Rico (1) |
| Highest Degree of Education Achieved | Secondary (5)  
                              | Undergraduate (20)  
                              | Master (3) |

Table 6. Background information of participants who met the clicking criteria.
All participants were paid $11 for their time for a session that lasted approximately 20 minutes.

4.4.2 Materials and Design

*Auditory sentences with ser and estar*

A total of 18 target items were prepared. Each item had three versions of predication: (1) with *ser*, (2) with *estar* followed by a natural process, and (3) *estar* with an artificial process. All target sentences shared the structure in (161):

\[(161) \ [\text{NP}] \ [\text{Copula}] \ [\text{Intervening Material}] \ [\text{Color Expression}] \ [\text{Disambiguation}]\]

Participants heard the copula earlier in the target sentence, before they heard the color expression and the disambiguation phrase that provided the critical expression that helped them identify the correct picture. The intervening expression was inserted in order to create a window to observe any anticipatory process after the copula.

Example (162) shows an example of a target sentence in the three conditions:

\[(162) \ a. \ [\text{NP Las uñas}] \ [\text{Copson}] \ [\text{IM como puedes ver}] \ [\text{Col negras}].\]

‘The nails *are*SER, as you can see, black’

\b. \ [\text{NP Las uñas}] \ [\text{Cop están}] \ [\text{IM como puedes ver}] \ [\text{Col negras}] \ [\text{Dis de suciedad}].\]

‘The nails *are*ESTAR, as you can see, black with dirt’
In (162), the NP las uñas ‘the fingernails’ is the NP subject of the predication, and negro ‘black’ the color expression. In the ser condition (162a), the Disambiguation expression was omitted with the assumption that the picture of the permanent color would be identified with ser by default. In the estar with natural process condition (162b) the disambiguation phrase provided the information about a natural process—dirty fingernails. And in the case of estar with an artificial process (162c), the disambiguation expression revealed an artificial process - the fingernails have been painted.

The 18 critical sentences were grouped into 3 item groups and cycled to create 3 different presentation lists (i.e. Latin Square design). Each list contained 6 ser sentences, 6 estar with natural process and 6 estar with artificial process, such that each item group would appear in all three conditions across the three lists. A complete list of target items can be found in Appendix A.

In addition to the 18 target sentences, 36 filler sentences are also created for the experiment. These filler items had a similar configuration to the target ones; a sentence that contains either the ser or estar copula, as well as a color adjective, accompanied by a display of four pictures, one in each quadrant, with a small black cross centered in the middle of the screen. The copula in these filler items were either estar with a locative expression or identificational uses of ser.
These two types of copular uses were chosen such that participants would not be able to identify the target items by the presence of the copula. Additionally, they also include a color expression in a noun modifier position, such as *la mochila naranja* ‘the orange backpack’. In target items, the color expressions are the predicate of the copula, as in *el pelo es blanco* ‘the hair is white’. The color expressions are, therefore, syntactically distinct in target and filler items. The use of color expressions in both the target and the filler items hinted a fake purpose of the experiment testing color perception of various objects. In the case of the visual displays for filler items, they all included one target picture and three distractors. The following examples show a filler item of the locative *estar* category (6a), and the identificational *ser* (163b) one:

\[(163)\]

a. La mochila está, como puedes ver, en el perchero grande naranja.

   ‘The bookbag is *estar*, as you can see, in the big orange coat rack.’

b. Mi regalo es, evidentemente, el coche de bomberos rojo.

   ‘Mi gift is *ser*, evidently, the red fire engine.’

The same filler items are randomized with the target items in each of the three presentation lists. All audio files (18 x 3 + 36 = 90 sentences) are recorded with a female native speaker of Spanish (age: 28, from Mexico).

*Visual input*

A total of 54 displays were prepared (18 target and 36 filler bitmap files). Each target display was combined with the three corresponding target sentences. Each target display included four
images. One image showed an object where the color of the predication is inherent; another where that color is temporary but due to an artificial process; the third picture showed an object where the color is temporary but due to a natural process and the last picture shows an object with a different color. The images were selected from online sources and were controlled for saliency and ease of object recognition. The complexity of four images was also controlled within each display. An example image array for the target sentences in (162) is shown in Figure 4:

![Figure 4](image)

Figure 4. Top left corner is the *estar* + natural process condition (dirty fingernail), top right is the *estar* + artificial process condition (painted fingernails), bottom left is the *ser* condition (gorilla hands) and bottom right is the distractor.

In the case of Figure 4, the *ser* condition is depicted by the hand of a gorilla, which has naturally black fingernails, the *estar* with natural process condition by dirty fingernails and *estar* with an artificial process with fingernails that have been painted black. The distractor image shows fingernails that are not black. Participants were asked to listen to the whole sentence and click on the picture that best represents what they heard. The same set of displays were presented across the three presentation lists. Between the trials, the small black cross remained in the center of the
screen. Subjects were asked to fix their gaze on the cross in between trials for checking the calibration status.

The image layout in the case of filler sentences was similar to that of the target sentences: they included four images, all of them showing the color used in the sentence. Only one of the images displayed the object mentioned in the sentence, while the rest represented different objects.

Procedure

Participants were seated approximately 60 cm away from the monitor of a Tobii 1750 eye-tracking system. Their eyes were calibrated using the ClearView 5-point calibration software. Following a successful calibration, participants listened to the sentences in Spanish, and clicked on the object that matched the referential expression after each question. Their eye locations were continuously recorded at 50 Hz. Participants were randomly assigned to one of the three presentation lists. On each trial, a fixation cross appeared in the center of the screen. After the cross was fixated for one second, a slide with four images appeared. The target sentence started 500ms after the onset of the visual input. When the participant clicked on an object after the sentence ended, the monitor switched to the fixation cross.

4.4.3 Predictions

According to the framework for estar predications with open-scale adjectives presented in Chapter 3, I first predicted that there would be anticipatory looks towards the picture
representing the natural process as soon as the copula *estar* is processed in both conditions that include *estar* (e.g., (162b) and (162c)).

In the case of the *ser* condition, in the absence of any constraints on the prediction beyond the non-mutability of the predication, I predict that listeners might not make any anticipatory eye movements towards any of the pictures when hearing the copula, since all objects depicted have an inherent color that could appear in a *ser* predication.

### 4.5 Results

#### 4.5.1 Results from the entire group of native listeners

The graphs below show the changes in the looks to each visual candidate (permanent, natural process, artificial process and distractor) as a function of time. In these graphs, the fixation proportion functions were realigned at the onset of each of the five critical windows: Subject, Copula, Intervening Material, Color and Disambiguation. Since target auditory files had various duration for each window, the duration assigned to each of those 5 sections is calculated by averaging over the duration of the target sentences.

The fixation rate function is generated by first converting the x- and y-coordinates of the fixation for each sample to the meaningful Area of Interest (AOI) type: such as permanent, temporary with natural process and temporary with artificial process. Since sampling rate was 50Hz, this measurement is taken every 20ms. At each point in time, the subject might be looking at one of the four pictures on the display: For instance, in the fingernail example provided above, the
participant may have looked at either the gorilla fingernails, the dirty fingernails, the painted fingernails or the non-black fingernails at each point where gaze data is collected. Since each participant heard a total of six trials per condition, data is collected for sentences including the *ser* copula up to six times. For each time point, there is a total of up to 6 different entries per participant, each entry representing the picture he or she is looking at on that particular point. Those total values are then turned into ratios for each condition. For instance, during each of the 6 *ser* trials he hears, at the same point during the sentence, a particular subject might be looking four times to the picture representing the *ser* condition, once to the *estar* with artificial process and once to the distractor. In this case, the ratios for that subject at that time point would be 4/6 to *ser*, 1/6 to *estar* with artificial process, 1/6 distractor and 0/6 *estar* with natural process. These ratios were averaged across all included participants to generate the group mean functions in the graphs. Following the convention of the visual world paradigm, fixation data from only correctly responded trails were included for generating the fixation proportion functions. A mixed-effects logistic regression analysis on the proportion of looks towards the picture representing the natural process was carried out⁶, with the proportion of looks towards the picture with the natural process as the dependent measure and the condition of the sentence—*ser*, *estar* with natural process and *estar* with artificial process—as the fixed factors. The *ser* condition was used as the reference level for the predictor factor.

⁶ `lmer(elogit for target pic ~ Condition+(1|Subj)+(1|Item), data=senh)`
**Ser Condition**

The gaze data in the *ser* condition shows no preference for any of the four pictures up until the point where the sentence is disambiguated, when they start directing a higher number of looks towards the picture representing the permanent condition. As can be seen in Figure 5, the graph lines representing the four conditions associated with the pictures in the display show these images were all receiving a similar number of looks throughout the sentence, and only during the final disambiguation stage, the picture representing the permanent color starts to receive a higher number of looks.

Let us now consider the data collected for the sentences that included *estar* with a natural process, as seen in Figure 6:
In Figure 6, looks towards the picture representing the natural process started increasing immediately after the copula. The looks to the natural process picture were maintained until the end of the sentence, where it gets confirmed in the disambiguation area.
The fixation pattern in the earlier window of Figure 7 are similar to those of Figure 6 (the *estar* with natural process). Once again, as soon as they heard *estar*, the participants started looking towards the picture representing the natural process. Only after an artificial process unfolded, participants began to direct more looks towards the picture depicting an artificial process.

A sentence window by window analysis shows the differences between the 3 different conditions more clearly. In the first window, SUBJ, where the participants hear the subject of the predication, the proportion of looks to all images is approximately equal in all three conditions, as Figure 8 shows.

![Figure 8. Window 1 Fixations](image)

The second window presented the copula. During this window, the three conditions start to show differences in the proportion of looks. In the *ser* condition, participants show no differences in their ratio of looks to any of the four pictures. In the *estar* conditions, both the natural process and artificial one, participants start to show an increased number of looks towards the picture representing the natural process. Figure 9 shows this.
The third window presented intervening material between the copula and the color expression. At this point in the sentence, participants have already heard the copula, but not the color expression, which makes this the critical window for observing anticipatory effects. If *estar* projects an expectation of a predicate with a natural process, it is at this point of the sentence where that expectation should start to show a difference in the fixation patterns. Figure 10 shows the proportion of looks in the third window.
In this third window, the *ser* condition maintains the same pattern that previous windows did. All pictures received a very similar proportion of looks. In the case of *estar*, both in the natural and artificial process conditions, participants maintained their larger proportion of looks towards the picture with the natural process. For the statistical analysis, a set of mixed-effects logistic regression models using R software and the lme4 package was constructed. This model tested the contribution of the condition of the sentence towards the likelihood of the looks to the natural process picture. The dependent measure was the log ratio of looks to the natural picture and the total number of fixations coded within the window. Since the core question was whether the two *estar* conditions induced anticipatory fixations, the *ser* condition was chosen as the reference condition. The full model expressed here did not converge, so the random effect slopes were deleted from the model and the random slopes were specified as (1|Subj) + (1|Item). Table 7 shows the results of this analysis.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>SE</td>
<td>t</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>-5.634e-03</td>
<td>8.859e-04</td>
<td>-6.360</td>
</tr>
<tr>
<td>Estar natural</td>
<td>4.939e-06</td>
<td>1.220e-03</td>
<td>0.004</td>
</tr>
<tr>
<td>Estar artificial</td>
<td>-1.172e-03</td>
<td>1.231e-03</td>
<td>-0.952</td>
</tr>
</tbody>
</table>

N = 522 Total number of observations. ***p < .001, **p < .01, *p < .05, †p < .1

Table 7. Summary of mixed effect regressions on the fixation ratios. All participants.

---

7 lmer(elogit for target pic ~ Condition+(1+Condition|Subj)+(1+Condition |Item), data=senh)
The results of the mixed effect models revealed no difference between the looks to the picture depicting the natural process in any of the three windows.

A pairwise analysis of the data collected during this third window was also carried out, using the R software and the lsmeans package\(^8\). This analysis compared the log ratio of looks to the natural picture and the total number of fixations detected within each window. This analysis finds no significant contrast between the ratio of looks towards the natural process picture in sentences where *estar* was heard and disambiguated to a natural process and those where *estar* was heard but it disambiguated to an artificial process conditions \((z = 0.295, p = 0.9533)\). This result is not surprising, given that during the third window, both the natural and artificial process sentences have provided the same information: the subject, the copula *estar* and the intervening material.

In the case of the comparison between the *ser* and *ser* and *estar* with natural process conditions, a statistically significant difference in the proportion of looks towards the picture of the natural process and the rest was found \((z = -2.799, p = 0.0142)\), whereas a marginal difference was detected in the case of *ser* and *estar* with artificial process conditions \((z = -2.256, p = 0.0622)\). These two last results show that hearing the copula *estar* leads to a significant number of looks towards the natural process picture, which leads to the conclusion that native speakers are predicting the presence of such a natural process based only on the processing of the copula *estar*.

\(^8\) lsmeans(regression model, pairwise~Condition)

205
While the anticipatory processing with the copula *estar* was confirmed for the entire group, a great amount of individual variance in the eye-tracking data was also observed. In order to further investigate the possibility of individual variation and what possible factors could be behind it, three subgroups within the participants were established. The next section details the criteria used in determining the groups and what the analysis of the data from those subgroups revealed.

### 4.5.2 Subgroups of Participants

In order to divide the group of 30 participants into three subgroups, the criteria that was followed was their overall accuracy in the clicking task. The rationale behind this is that it might be that case that subjects that perform better at the clicking accuracy task also show differences in their incremental parsing of the sentences. The overall accuracy of all participants was examined and the three groups, a Higher Accuracy group with 9 members, a Middle Accuracy Group with 12 members and a Lower Accuracy Group with 9 members, were established. The graph in Figure 11 shows a histogram of participants according to their overall accuracy on the clicking task.
The background data from the questionnaire for each of the three groups are shown in Table 8:

<table>
<thead>
<tr>
<th></th>
<th>Higher Accuracy Group</th>
<th>Middle Accuracy Group</th>
<th>Lower Accuracy Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy Range</strong></td>
<td>88.8% to 94.4%</td>
<td>83.3%</td>
<td>55.5% to 77.7%</td>
</tr>
<tr>
<td><strong>Overall Accuracy</strong></td>
<td>91.3%</td>
<td>83.3%</td>
<td>70.6%</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>29.4</td>
<td>28.7</td>
<td>29.4</td>
</tr>
<tr>
<td><strong>Hours Spanish per Day</strong></td>
<td>5.4</td>
<td>5.1</td>
<td>5.8</td>
</tr>
<tr>
<td><strong>Months in US</strong></td>
<td>26.5</td>
<td>32.2</td>
<td>13.1</td>
</tr>
<tr>
<td><strong>Dialect</strong></td>
<td>Spain(4)</td>
<td>Spain(4)</td>
<td>Spain(5)</td>
</tr>
<tr>
<td></td>
<td>Mexico(3)</td>
<td>Mexico(4)</td>
<td>Mexico(3)</td>
</tr>
<tr>
<td></td>
<td>Venezuela(1)</td>
<td>Puerto Rico(1)</td>
<td>Bolivia(1)</td>
</tr>
<tr>
<td></td>
<td>Bolivia(1)</td>
<td>Colombia(1)</td>
<td></td>
</tr>
<tr>
<td><strong>Higher Degree</strong></td>
<td>Undergrad (7)</td>
<td>Secondary(5)</td>
<td>Secondary (2)</td>
</tr>
<tr>
<td><strong>Education Achieved</strong></td>
<td>Master (2)</td>
<td>Undergrad (5)</td>
<td>Undergrad (7)</td>
</tr>
<tr>
<td><strong>Reaction Time</strong></td>
<td>8347.93 ms (SD = 1686.5)</td>
<td>8578.03 ms (SD = 1765.9)</td>
<td>8920.90 ms (SD = 1483.9)</td>
</tr>
</tbody>
</table>

Table 8. Background Information of Participants per Subgroup
Several conclusions can be reached by looking at the information contained in Table 6. Age and hours of use of Spanish per day do not seem to be a factor in performance in the experiment, as all groups have comparable age and hours per day functioning in Spanish averages. The amount of time that participants have spent in the US since they left their Spanish speaking countries appears perhaps to be counter-intuitive, as those participants that have been, on average, living in the US longer perform better than those that moved more recently. In an extreme case, a member of the Higher Accuracy group was born in the US and has spent the majority of her life in an English-speaking country. The dialectal variation within each subgroup also seems to yield no apparent differences, as both dialects that supply the highest overall number of participants in the experiment—Spanish and Mexican—are also the two dialects that have the most representatives in both subgroups. Although there are participants from other dialectal backgrounds in the subgroups, their numbers are too small to allow us to make any claim about their influence. Nevertheless, both dialectal and educational background are factors that could prove significant, if a larger number of participants were available, and as such, present an opportunity for future avenues of investigation. An additional factor considered was the average Reaction Time of members of the group. This Reaction Time is a measurement of the time, in ms, between the beginning of the auditory input and the time when the participant clicks on a picture. The inclusion of this measurement is intended to show whether subjects that perform better at the clicking task also click quicker than their lower performing counterparts. In any event, although the Reaction Time increases as accuracy decreases, the values fall within the standard deviation, so, at this point, this factor also reveals no differences between groups.
Eye-Tracking Data

After establishing the groups and trying to ascertain any possible factors affecting their performance from their background information, the gaze data from each subgroup was analyzed. In the first place, their data for the ser condition was examined. Figure 12 shows the comparison between the Higher, Middle and Lower Accuracy groups.

The graphs in Figure 12 show the gaze information for the Higher Accuracy group on the left, Middle Group in the center and the Lower Accuracy group on the right. As it was the case in the collective analysis, in the ser condition, participants do not make any anticipatory looks. In both cases, the graph shows that subjects need to wait until the final part of the sentence in order to determine which picture best represents the sentence they are hearing.

Next, the gaze data for the estar with a natural process was examined. The resulting graphics can be seen in Figure 13.
These same results are repeated in the *estar* with the artificial process condition, as can be seen in Figure 14.

Table 9 summarizes the results of the regression analysis of the fixation ratios during the third window of the sentences. The same software and method used for the analysis of the entire group of participants was used in the analysis of the three sub-groups. The third window is the one immediately after the copula is heard, and it is the point where anticipatory looks can be observed. In the analysis, the log ratio between the looks towards the picture depicting the natural process to all the fixations in this third window was examined, in cases when the subject heard the copula *ser*, represented by Condition TS (Intercept of the model) and in those cases when they heard *estar*, represented by Conditions TE (*estar* with a natural process) and TG (*estar* with artificial process). The rationale behind this choice is that during this third window, the only information that the sentence has provided is the subject of the predication, the copula
and the intervening material. Participants showing a significantly higher proportion of looks

towards the picture depicting the natural process could be interpreted as a sign that the
requirement of a natural process for the interpretation of *estar* with color predications has been
incorporated into the copula itself, and at the point of the third window are already predicting the
presence of said natural process. Table 8 shows the data for the High, Middle and Lower
Accuracy groups.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>High Accuracy Group</th>
<th>Middle Accuracy Group</th>
<th>Lower Accuracy Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>W3</td>
<td>W3</td>
<td>W3</td>
</tr>
<tr>
<td></td>
<td>Est.</td>
<td>SE</td>
<td>t</td>
</tr>
<tr>
<td>TS (Intercept)</td>
<td>-0.8257</td>
<td>0.2219</td>
<td>-3.722</td>
</tr>
<tr>
<td>TE</td>
<td>1.2307</td>
<td>0.2952</td>
<td>4.169*</td>
</tr>
<tr>
<td>TG</td>
<td>1.1836</td>
<td>0.2962</td>
<td>3.995*</td>
</tr>
</tbody>
</table>

Table 9: Summary of mixed-effect regressions on the fixation ratios.

Table 9 shows an effect for the High Accuracy group for the TE and TG conditions, indicating a
significant higher number of looks towards the picture representing the natural process when the
copula *estar* was heard.

A pairwise analysis of the three conditions based on the data produced by the regression model
was conducted for all three sub-groups. The same method as in the case of the entire group of
participants was used. For the Higher Accuracy Group, the analysis reveals a significant
difference between the TS-TE conditions and the TS-TG conditions. In the case of the Middle
Accuracy Group, a significant effect was found for the TS-TG conditions and a marginal one for
the TS-TE condition. In the case of the Lower Accuracy Group, the analysis reveals no significant effect. Table 10 summarizes these results.

<table>
<thead>
<tr>
<th>Pair</th>
<th>High Accuracy Group</th>
<th>Middle Accuracy Group</th>
<th>Lower Accuracy Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>z</td>
<td>p</td>
<td>z</td>
</tr>
<tr>
<td>TE-TG</td>
<td>0.166</td>
<td>0.9849</td>
<td>-0.448</td>
</tr>
<tr>
<td>TS-TE</td>
<td>-4.169</td>
<td>0.0001</td>
<td>-2.171</td>
</tr>
<tr>
<td>TS-TG</td>
<td>-3.995</td>
<td>0.0002</td>
<td>-2.581</td>
</tr>
</tbody>
</table>

N = 170 High, N = 151 Middle, N = 120: total number of observations.

Table 10. Summary of pairwise analysis on the fixation ratios.

The observation of the differences in the gaze patterns in both estar conditions between the Higher, Middle and Lower Accuracy groups is revealing of one important difference. The hypothesis that the presence of a natural process in estar with color predications in Spanish is affecting the way sentences are processed seems to hold in the case of certain individuals, but not others. Since the participants were divided in groups according to the accuracy of their clicking responses, the relationship between clicking accuracy and the looks to the picture representing the natural process was examined. Figure 15 shows the relationship between accuracy and the log of the ratio between looks to the natural process picture and the distractor picture in the third window of the sentence. The two graphs represent the accuracy in the estar with natural process (TE) and artificial process (TG) conditions.
High TE N = 50, TG N = 57; Mid TE N = 36, TG N = 41; Low TE = 39, TG = 32: Total number of participants

Figure 15. Accuracy in Clicking Task and log ratio between looks to natural process picture and distractor

The y axis of the graphs represents the log of the ratio between the looks directed towards the picture depicting the natural process and the looks to the distractor picture. The x axis represents the percentage of correct responses on the clicking task that the participant had in each condition. Each data point in the graph represents the fixation likelihood towards the natural process picture during the third window for each item where the participant clicked on the correct picture at the end of the sentence. For example, if a participant had an 83.3% accuracy in the TE condition, i.e. when the sentence contained the *estar* copula and disambiguated to a natural process predication, this means he clicked correctly on 5 of the 6 total TE trials. The relevant proportion of looks for those 5 items is retrieved from the eye-tracking data and are then plotted into the graph.
Pearson's product moment correlation was used to test the correlation between the log ratio of looks to the natural picture and the distractor and the accuracy in the clicking task data\(^9\). In the *estar* with artificial process condition (TG), the association between the log ratio of looks and the accuracy is significant.

The same analysis was carried out on the data of each of the three group of participants separately. For the Middle Accuracy Group, a correlation effect was found in the *estar* with artificial process condition but no other effect was found. Table 11 shows the results of the analysis.

<table>
<thead>
<tr>
<th>Condition</th>
<th>All Groups</th>
<th>High Accuracy</th>
<th>Middle Accuracy</th>
<th>Lower Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TE</strong></td>
<td>r = 0.01254484</td>
<td>r = 0.2214519</td>
<td>r = -0.1682917</td>
<td>r = -0.0111485</td>
</tr>
<tr>
<td><strong>TG</strong></td>
<td>r = 0.1773961**</td>
<td>r = 0.2218944</td>
<td>r = 0.3316066*</td>
<td>r = 0.0117968</td>
</tr>
</tbody>
</table>

High TE N= 50, TG N = 57; Mid TE N = 36, TG N = 41; Low TE = 39, TG = 32; Total number of participants **p < .01, *p < .05, †p < .1

Table 11. Summary of correlation between log ratio looks to the natural picture and the distractor and the accuracy in the clicking task data.

The differences between the results in both *estar* conditions can be attributed to a stronger lexical disambiguation that the artificial process condition had, which lead to a higher overall accuracy (86.8%) than the *estar* with natural process (82.6%). Before a final analysis is provided in Section 4.6, the data gathered in the clicking accuracy task is examined in the following section.

\(^9\) cor.test (percentage, elog, method = "pearson")
4.5.3 Clicking Accuracy

The analysis of the clicking accuracy data gathered during the experiment can also be used to gain some insight into the interpretation of copulas in Spanish. The clicking task that participants were given involved clicking on the expected picture after listening to the sentence accompanying each of the target slides. According to the established criterion, in order for a subject’s data to be considered valid for the experiment, they need to click correctly in at least half of the trials per condition, i.e. at least three correct each in the *ser*, *estar* with natural process and *estar* with artificial process. Table 12 presents the clicking accuracy data per condition for the 28 subjects that met that criteria.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Total</em></td>
<td>80.3%</td>
</tr>
<tr>
<td><em>Ser</em></td>
<td>72.2%</td>
</tr>
<tr>
<td><em>Estar + natural process</em></td>
<td>82.6%</td>
</tr>
<tr>
<td><em>Estar + artificial process</em></td>
<td>86.8%</td>
</tr>
</tbody>
</table>

Table 12. Clicking Accuracy.

The data in Table 12 shows that the overall accuracy of subjects that met the clicking criterion is 80.3%. As far as accuracy per condition, the *estar* with artificial process condition is the most accurate, followed by the *estar* with natural process, and finally, the *ser* condition. This distribution is not surprising, given that only the items in the *estar* conditions have an explicit disambiguation at the end of the sentence, while the *ser* condition does not. The percentages themselves, although they might appear low for native speakers of Spanish, are not entirely
surprising when compared with the ones found in Holtheuer et al. (2011) as given in Table 12, repeated here.

![Table 13. Experiment 1 results from Holtheuer et al. (2011)](image)

As Table 13 shows, in the case of adult native speakers, in cases where there is a competition between a possible *ser* and *estar* interpretation, represented in the Same Character condition in Table 11, the proportion of *ser* and *estar* in their elicitation task is 75 to 24. Although the data from Holtheuer et al. (2011) is from an offline interpretation of the copula, it shows that, even in the case of adult native speakers, the copula system in open-scale adjective predication is an unstable one. In order to obtain a deeper understanding of the clicking patterns, the type of errors that were committed within each condition are also analyzed. The results are summarized in Table 14.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Type of Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ser</td>
</tr>
<tr>
<td><strong>Ser</strong></td>
<td>-</td>
</tr>
<tr>
<td><strong>Estar + Natural</strong></td>
<td>56%</td>
</tr>
<tr>
<td><strong>Estar + Artificial</strong></td>
<td>63.2%</td>
</tr>
</tbody>
</table>

Table 14. Accuracy Task Errors per Condition.
The data in Table 14 shows that in the case of ser and estar with a natural process, the errors are more or less evenly distributed between the other two conditions, which suggests that subjects might be unsure on which picture to click. The results for the estar with an artificial process, on the other hand, reveal a preference for the ser condition. This pattern can be attributed to the distribution that the corpus analysis reflected: an artificial process does not usually get represented with an estar copula, and thus if the picture is interpreted as such, it should be harder to select it if the copula estar was heard in the sentence.

4.6 Conclusions

Two main conclusions can be extracted from the results of the eye-tracking experiment. On the one hand, the gaze data suggests that certain native speakers of Spanish do seem to have an anticipatory effect in the online processing of estar structures with color expressions. The intuition that this type of estar predication only applies in cases where the change in color is the result of a natural process affecting the individual subject of the predication can be seen in the eye movements that speakers directed to the pictures depicting such processes upon hearing the copula, and before the sentence disambiguated completely. The ser copula, on the other hand, does not seem to have any anticipatory effect during the parsing of the sentence.

Whether the effects observed for estar carries over to any open-scale adjective or it is restricted to color adjectives still remains to be tested, although the results of the experiment are promising. Furthermore, these anticipatory looks are only observed for some individuals—in the case of this experiment, identified by performing better in the task that required them to click on the correct
picture after hearing the sentence. Determining the exact factors that govern this distinction among individuals is still a question for future research endeavors, but for the purposes of this experimental approach, the effect that the natural process has on the online processing of sentences containing *estar* with color predications has been attested.

Additionally, the data extracted from the clicking task also reveals valuable information about the interpretation and difficulties associated with the process that is associated with *estar* structures with an open scale predication. As the semantic model predicts, the pragmatically-determined process seems indispensable to the validation of the presence of *estar*, and this constraint puts a further load on the speaker’s part, since they need to be able to recognize the situation as one that includes such a process. Beyond the nature of *ser* and *estar* requiring that the quality associated to the subject of the predication displays certain properties of mutability, the necessity to recognize the process, in the case of open-scale adjectives, as one further requisite can create divergent interpretations and copula distributions. One only needs to consider how, in some of the target items previously presented, not identifying a certain river as being naturally red, or as being green because of a particular celebration leads to some unexpected copula choices.

As a general analysis, the results of the experiment not only confirm that the process associated with certain *estar* structures plays a role, in certain individuals, on the online interpretation of said structures, but also plays a central role on the offline one, particularly in the case of copula distribution. It seems that further investigations along the lines of these processes on open-scale
predications could be fruitful not only from the theoretical and experimental point of view, but also in the identification of a sociolinguistic factor that could be affecting the online processing of copulas in Spanish.
Chapter 5: Conclusions

5.1 Summary

The present study set out to provide an account for the distribution of *ser* and *estar* in Spanish in adjectival contexts, articulated around the scalar properties of that type of predication. Based on the observation of the data present in Spanish and the stronger and weaker points of the vast literature available on the topic, three guiding principles were formulated:

i) The distribution of *ser* and *estar* is not dependent on temporal properties of the predicates, but rather on the scalar nature of the predicate that combines with it. *Ser* requires that the predication describe a single degree along the adjectival scale, while *estar* requires that the predication describe an interval on that scale.

ii) The distribution non-evidential *estar* is sensitive to the scalar structure of the adjective present in the predication. Open scale adjectives require, on top of defining a mutable property of the subject, a natural process that drives that change. Closed scale adjectives only require the mutable property. These properties are not part of the lexical information provided by the adjective, but rather are provided by the comparison class necessary to compute the truth condition of the predication.

iii) Evidential uses of *estar* need to be considered when constructing a theoretical framework for copula distribution in Spanish, since they have their own idiosyncratic behavior. These uses of *estar* are also characterized by describing an interval along a
scale, although their intervals are nor predicated on the mutable property of the predication, but rather on the distance towards the standard of comparison. Although evidential uses of estar have a distinct interpretation from the non-evidential uses, their basic mechanism can be accounted in terms of the interval requirement.

The first two considerations stand behind the first main theoretical decision that is adopted for the proposal; the present study is articulated around the scalar framework based on Kennedy and McNally’s (1999, 2005) influential proposal. The reason behind it is twofold: on the one hand, in order to incorporate the importance of the scalar structure of adjectives into the proposal, basing it on classical interpretation of scales provides not only a well-defined and powerful tool with which to build, but also a wealth literature that can provide ideas and approaches that might have been overlooked. A second reason that underlies the choice of a scalar framework for the present study is how it can provide a mechanism to account for the first consideration: a scalar account needs not the be constructed around temporal considerations, and, as such, it can easily provide a mechanism that formalizes the distribution of ser and estar where the concept of permanent and temporary is a consequence of the proposal, and not a necessity.

The third consideration deals with the importance of incorporating evidential uses of estar, which have been traditionally either ignored or else present a problem for traditional accounts based around temporal features of the predications. These evidential uses of estar do not share the same features than the non-evidential uses, in the sense that they do not require that the adjectival expression in the predication is a mutable property of the subject. The theoretical
framework presented for the non-evidential proves flexible enough to support even cases where not only temporal features are irrelevant, but that analysis can be taken to remove even the potential for change that drives their distribution.

Once the theoretical framework was set, the basic claim was laid out: *Ser* and *estar* are copulas that only differ in the type of predication that they can appear with, and formalizing those different types of predications is the principal task of any attempt to account for their distribution. The main claim then became that predicates that appear in *ser* structures are characterized by involving one single degree along a scale, while predicates that appear in *estar* structures are characterized by involving two degrees along a scale. This very basic claim allows us to capture all *ser* contexts, as well as both the non-evidential and evidential uses of *estar*. Moreover, it also allows for the capturing of the differences observed between the distribution of *ser* and *estar* in closed scale and open scale adjectival contexts. This distinction is based on one of the main innovations suggested in the study: the presence of a natural process that needs to be present in *estar* predications with open scale adjectives.

The experimental portion of the present research centers around providing evidence for any cognitive content associated with the natural process and set a baseline for future research along the lines of the investigation of the online processing of copulas in Spanish. In this respect, the experiment from Chapter 4 has shown that adult native speakers of Spanish make early decisions concerning the presence of a natural process in the context of open scale adjectives, and more specifically, color expressions. The analysis of the eyetracking data shows that native speakers
predict the presence of this process at the point where they hear the copula *estar*, and similarly, predict the presence of a non-mutable property in the predicate of *ser* structures.

### 5.2 Limitations and Strengths

#### 5.2.1 Relevance

The relevance of the present study can be determined on two fronts. On the one hand, it contributes a new perspective to the existing literature of theoretical approaches to the distribution of the *ser* and *estar* copulas in Spanish. Building on existing theoretical frameworks, it provides several insights, based on observed data, that have so far gone unnoticed: the way in which the scalar structure of adjectival expressions affects copula choice, and how closed scales and open scales require slightly different conditions in order to give rise to either copula. The importance of the scalar structure of adjectives in copula distribution is not only one of the main theoretical and empirical contributions of the present study, but also one that can have the more far reaching effects. The relevance of the process associated to these *estar* predications could be implemented in second language acquisition situations, especially language learning.

#### 5.2.3 Limitations

There are also several limitations to the present study. The main one is the circumscription of the predicates investigated only to those including an adjectival expression. While this limitation in the scope of predications under study can be justified in terms of adjectives providing the best variety of situation in order to establish a baseline for the distribution—adjectival contexts
provide *ser* exclusive, *estar* exclusive and alternating contexts—extending the present analysis to other types of predications can present challenges.

Another area in which the present analysis of the copulas might need to be extended in future extensions is the variety of Spanish used throughout the study. Using a similar rationale to the one used in the choosing of adjectives as the starting point for the analysis, the standard variety of Spanish was chosen as the dialect of choice. There is, however, a certain degree of variation in copula use in Spanish (Silva-Corvalán, 1986; Geeslin & Guijarro-Fuentes, 2008). As is the case in the limitations imposed by restricting the context of predicates only to adjectives, restricting the analysis to one dialect, predominant as it may be, leaves room for cases where dialectal variation requiring that the theoretical framework be modified or extended.

In the case of the experimental portion of the research, the fact that to generate an open scale context for the target sentences only color expressions are used can result in the anticipatory effect suggested by the eyetracking data be partially due to the interpretation of color expressions themselves. In order to strengthen the claim that adult native speakers of Spanish make early predictions regarding copulas in such contexts, other non-color, open scales might be necessary.

### 5.3 Future Research

Regarding the research put forward in this study, both in the case of the theoretical aspects and the experimental ones, some avenues present themselves for future research already present themselves at this point. The theoretical framework based on scalar properties of predicates
presented for adjectives can be extended to cover all contexts. Moreover, the experimental portion also presents several areas where future research can extend the results presented here into new areas of interest. The following sections provide some of the possibilities for future research.

5.3.1 Locative Expressions

Locative expressions and their almost exclusive distribution with *estar* represent one of the bigger challenges for any theoretical account of the copulas in Spanish. As far as the one presented here, locative expressions can be integrated unto the general paradigm of *estar* requiring a predicate that depicts an interval on a scale if it is assumed that all locative expressions use a scale of physical distance and that they always indicate separation between the subject of the predication and the position it occupies in space. A similar idea is proposed by Mangialavori (2013), when she states that “the *be at*—corresponding to the locative copula—maps a thing into a spatial frame and asserts that the thing occupies a specific point within this space. […] nonetheless, in being articulated by a locative verb, this situation will be essentially characterized in terms of a further object.” In other words, the conceptual interpretation is that locative expressions determine the physical position of an object in relation to another, different object. Extending this analysis, if we assume a generic scale along a dimension of physical distance, a locative expression would require that an individual projects a certain degree along that scale, and furthermore, that degree would be evaluated against another referential degree—thus creating an interval on that scale, and therefore becoming a predicate that would require the *estar* copula in Spanish. This implementation would remove the temporary vs permanent
location distinctions that create much of the confusion in the characterization of these locative expressions as estar exclusive. For example, given two cases of locative uses of estar, as in (164):


The Sun is in the center of the Solar System

‘The Sun is at the center of the Solar System.’

b. Pepe está en la biblioteca.

Pepe is in the library

‘Pepe is at the library’

Example (164a) is a case of a rather permanent situation, while (164b) is one that is trivially temporary in most contexts. Following traditional accounts for ser and estar, these examples are rather troublesome to explain. On the other hand, if we assume that, in both cases, the Sun and Pepe are individuals that project a degree of distance along a scale—degrees that in the given examples would coincide with the center of the Solar System and whatever library happens to be relevant in this context—and we establish their position as comparison with another referential point, in both cases a scalar interval is generated. There remains, naturally, the problem of accounting for this second referential point, what is its nature, how is it established and whether it is stable between contexts and individuals. Nevertheless, the line of reasoning here presented seems promising in the integration of locative expressions into the proposed estar paradigm.
Regarding the small number of locative expressions that appear with *ser*, one possible explanation has been offered by Marín (2004), who considers them not strictly locative expressions, but rather falling within the scope of the identificational uses of *ser*. See example (165):

(165) La salida es esa puerta.  
    The exit is *SER* that door
    ‘The exit is that door’

In the previous example, the presence of *ser* can be interpreted as not representing a strict locative predication, representing the physical location of the exit, but rather an identification of the exit with a particular door. In this manner, (165) would be an example along the lines of the one in (166).

(166) La salida es el lugar por donde se sale.  
    The exit is *SER* the place by whom SE exits.
    ‘The exit is the where you leave.’

In both cases, the *ser* copula, although it does reference a physical place, is closer to identifying the exit with that place, not telling us where it is. This conception can be clearly seen in cases where this identificational reading would not be available, *ser* becomes unacceptable. For instance, on the previous example (164b), Pepe can be said to be located at the library, but he
cannot be identified with the library. Accordingly, the following sentence is unacceptable, as in (167):

(167)  *Pepe es la biblioteca.
       Pepe  issER the library
       ‘Pepe is the library’

From the previous example, it seems clear that the locative uses with *ser only really work in cases where it is possible to identify the subject with the location involved. As such, they would fall under the *ser paradigm.

5.3.2 Progressive and Passive Voice

Contexts for verbal predication with copulas in Spanish includes two contexts: Progressive form and Passive Voice structures. As previously mentioned in the section dealing with the distribution of copulas, the Progressive form is always constructed with *estar (168a) and the Passive Voice with *ser (168b):

(168)  a. Mi amigo {*es/está} escribiendo una carta a sus padres.
        my friend {issER/iSERAR}writing a letter to his parents.
        ‘My friend is writing a letter to his parents.’

       b. Mi amigo {fue/*estuvo} detenido por la policía.
          my friend {wasSER/wasESTAR} arrested by the police.
‘My friend was arrested by the police.’

The progressive aspect is very often interpreted in terms of a stage of an event (Asher, 1992, for example). The framework proposed for *estar* and adjectival predications in the present predication can be extended to incorporate progressive by extending the theoretical framework presented to include a temporal component. In this context, a productive extension could be to refer to the homomorphic relationship between the scalar properties of verbs and the temporal progress of the events they denote (Dowty, 1991; Kennedy & Levin, 1999). This homomorphic relationship could be incorporated to the analysis of *estar* as requiring a predication that denotes two degrees on a scale by stipulating that, in the case of the progressive, that scale also include a time reference. In this manner, the predicate in example in (5a) could be characterized as incorporating two degrees at two different times; one where my friend begins to write a letter and one where he finishes it. The first point in time would also include a starting degree of writing a letter and the final one a degree where the letter is completed. Naturally, this basic idea needs to be formalized and carefully put to the test, but it seems nevertheless promising.

In order to construct Passive voice structures, as previously mentioned, Spanish uses the copula *ser* exclusively. The Passive voice in Spanish is assumed to determine a single event in time (Sansó, 2001). In this case, and applying the same principle as in the case of progressive structures, the single event determined by the Passive will be reflected as a single degree in the homomorphic scale associated with the event. As such, the Passive voice would fall under the paradigm given for *ser* in this dissertation.
5.3.3 Nominal Predications

The distribution of copulas with nominal predications in Spanish also lends support to the theoretical framework of the present study. Nominal predications are almost exclusively found in *ser* contexts. This distribution was difficult to explain in terms of traditional accounts, since the temporal permanency or not of the nominal predication played no apparent role in the choice of copula. See example (169):

\[(169) \quad \text{a. Pepe \{es/*está\} un ser humano.} \]
\[
\text{Pepe \{ issER/isESTAR \} a human being} \\
\text{‘Pepe is a human being.’} \\
\text{b. Pepe \{es/*está\} un profesor.} \\
\text{Pepe \{issER/isESTAR\} a teacher} \\
\text{‘Pepe is a teacher.’}
\]

From the examples in (169), it is clear that (169a) is a predicate that needs to be interpreted as permanent or an Individual Level, and (169b) is one that describes a stage in the life of Pepe. And yet, they both need to necessarily appear with the *ser* copula. These examples have been traditionally assigned an identification reading (Marín, 2004), that is to say, the copula in (169) is assigning to the individual subject of the predication the features associated with the nominal expression in the predication, thus the identificational reading.
Nouns are considered to project no scalar structure in the literature (Morzycki, 2009), and, as such, are underspecified in this respect. However, as previously mentioned, nominal expressions are also found in *estar* contexts, as in (170):

\begin{enumerate}
\item Pepe estuvo torero.
\item Pepe was\textsubscript{ESTAR} bullfighter
\item ‘Pepe was brave.’
\end{enumerate}

These nominal expressions are considered to be cases of gradable expressions that grant access to the properties associated with the nominal expression (Constantinescu, 2011; Beltrama & Bochnak, 2011), often expressed in terms of prototypes. In this manner, once the properties associated with the noun become available, there is a scalar structure which can provide the necessary interval in order to generate an *estar* valid predicate. These nominal predications with *estar* fall into the same parameters than the adjectival ones, in the sense that the noun provides the same semantic features than an adjective would. According to this interpretation, they can be included along the rest of the adjectival expressions, with the caveat that the scalar information is provided by a noun. Provided that this last characteristic is given a formal interpretation, examples such as those found in (170) should pose no further problems.

### 5.3.4 Dialectal Variation

The present study has focused on standard variety of Spanish. There is however, a great deal of variation regarding copula use in Spanish, with an extensive literature: variation due to language
contact (Geeslin & Guijarro-Fuentes, 2008), language acquisition (Aguilar-Sánchez, 2012; Cheng, 2002). One of the objective of this study is to provide a baseline of the interpretation of both *ser* and *estar* in Spanish, based in adjectival predicates, that could potentially be expanded to include all possible uses of the copulas. In this respect, dialectal variation can provide a richness of opportunities where the framework presented here can be tested and from where it can grow. Another limitation of the present proposal is the challenge presented by dialectal variation. Throughout this dissertation, certain judgments have been formulated about the acceptability or not of certain structures. These judgements represent the combination of several criteria: my own personal intuitions, as a native speaker of Spanish, materials from corpora and online sources, and judgments gathered from other speakers of Spanish. Naturally, a context as wide as copula choice in adjectival prediations, is going to give cause for discrepancies in the acceptability of certain structures. As I pointed out at the beginning of this dissertation, I do not intend for my judgements to represent an absolute guide of acceptability, but rather a baseline from where a theoretical framework can be formulated. Based on the different sources consulted in order to determine acceptability, there is an important observation that can be made about the uniformity in the choice if copulas throughout the dialectal continuum: variance in acceptability seems to grow as context-independent points along the adjectival scale are removed. In the case of closed-scale adjectives, with their salient endpoints on the scale, judgements are very stable in all dialects. In the case of open-scale adjectives, judgements on the acceptability of *estar*, which depends on the presence of a natural process, present a greater degree of acceptability. Some sources consulted, for example, consider events of painting or dying as determining a valid process for *estar*. Finally, in the case of evidential uses of *estar*, with neither endpoints nor
natural process requirements, display the greatest degree of dialectal variation. This type of variation in the use of copulas within the Spanish continuum which seems to correlate with salient endpoints on the scale agrees with the concept presented in this dissertation that Interpretive Economy is a factor in copula choice in Spanish. While dialectal variation is both interesting and challenging, the cases observed do not represent situations that fall outside the scope of the current theoretical proposal, but rather opportunities to expand its scope.

As just mentioned, certain dialects are very conservative in the types of predicates that can be used in evidential estar structures, while others—Mexican dialects, for instance—show a much wider range of options available. On a first approximation, evidential uses of estar where the predication involves a measure of the general quality of the subject seem to be universally observed in the Spanish-speaking continuum (examples from REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [online]. Corpus de referencia del español actual. <http://www.rae.es> [3/16/2017]

(171)  

<table>
<thead>
<tr>
<th>SPAIN</th>
<th>VENEZUELA</th>
<th>BOLIVIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Ya lo he probado y está genial.</td>
<td>'I have already tried and it is great.'</td>
<td></td>
</tr>
<tr>
<td>b. Ricardo está genial!</td>
<td>'Ricardo is great.'</td>
<td></td>
</tr>
<tr>
<td>c. [...] me invitó un puchero cochabambino que estuvo maravilloso.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
me invited. 3rd.SG a stew cochabambino that was wonderful

‘He invited me to a cochabambino stew which was wonderful.’

(172) a. El café estaba malísimo. Spain

the coffee was bad

‘The coffee was really bad.’

b. Este carro está malo Venezuela

tthis car is bad

‘This car is bad.’

c. Mi casa está terrible. Paraguay

my house was terrible

‘My house is terrible.’

In all cases in (171) and (172) the predication expresses the general quality of the subject, whether positive, as in (171), or negative, as in (172). And, in both sets of examples, the quality feature of the predication is such that the subject would not be able to change it, leaving only an evidential use reading available. These type of estar structures, as previously mentioned, are found all over the Spanish dialectal continuum. In addition, some dialects have extended these evidential uses to contexts that go beyond expressing the quality of the individual, and we can find examples such as: from REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [online]. Corpus de referencia del español actual. <http://www.rae.es> [3/16/2017]
The examples in (173) exhibit the same basic principle of the evidential uses of estar that were found in (172)—the subject cannot be expected to change the properties expressed in the predicate—with one difference; the predicate does not express the general quality of the subject, but rather a physical feature. These types of evidential uses of estar can be interpreted as extensions and innovations on the general case found in all Spanish dialects, although they can still be included under the umbrella of the general paradigm proposed in this study.

5.3.5 The Natural Process

Although the presence and importance of the process that is a necessary for an estar predication in open scale adjective predications is well motivated in the present study, both theoretically and empirically, much work still remains to be done in order to capture and formalize exactly how this process is generated and why some of them provide the necessary features for an estar
predication, while others do not. One possible research direction that could prove fruitful in this respect is trying to extend the analysis that Marín & McNally (2011) propose for the differences in the notions of inchoativity and telicity in reflexive psychological verbs in Spanish—aburrirse ‘to become bored’ and enfadarse ‘to become angry’, for example. In their analysis, they find that there is a correlation between psychological verbs projecting a left boundary on their event structure and the availability of their participles to appear in estar predications. Given that the participles of these reflexive psychological verbs all project open scales, it seems that these left boundaries could be an indicator of the presence of a natural process. This analysis could be extended beyond the realm of deverbal adjectives by a similar mean to the one previously mentioned for the inclusion of the progressive constructions; referencing the homomorphic relationship between the scalar properties of these verbs and the temporal progress of the events they denote (Dowty, 1991; Kennedy & Levin, 1999). While this proposed analysis could provide a reliable tool to predict which open scale adjectives and participles will be able to appear in an estar predication, it perhaps does not solve the ultimate question as to what exactly is the nature of these processes.

5.3.6 Experimental Research

One area where the experimental portion of the study could be extended is into the field of First and Second Language Acquisition. Acquisition of copulas has received plenty of attention, from the perspective of First Language (Sera, 1992; Schmitt et al, 2004), Second Language (Van Patten, 1985; Ryan & Lafford, 1992; Geeslin, 2003), Bilingualism (Silva-Corvalán & Montanari,
2008) and these are areas where the results of the eyetracking experiment could prove interesting.

A very straightforward future research opportunity involves trying to determine which, if any, anticipatory effect can be found in learners of Spanish in their online interpretation of copulas. Given that the process associated with open scale adjectives and *estar* is not explicitly taught anywhere—the present study being the first time it has been identified—it is not clear whether these learners would be able to identify this distribution of *estar*. Within this context of learners of Spanish, another area that could prove fruitful is not only whether the anticipatory effect found in native speakers is present, but also if explicitly teaching the distribution of *ser* and *estar* following the model presented here would affect the online interpretation of open scale adjectives and if learners would also develop that anticipatory effect as a result of explicit training.

In the case of bilingual and heritage speakers of Spanish, replicating the experiment with that type of subjects could also further our understanding on the nature of the process attached to open scale structures as predications of copulas. The current hypothesis is that those process are pragmatically determined and they are a consequence of extended exposure to these structures. In the case of bilinguals and heritage speakers, their exposure might not have proven enough to develop native-like intuitions. In any event, extending the research to those two groups of individuals seems like a natural avenue in order to develop a clearer picture of the nature and extent of the anticipatory effect found in native speakers. The same principle could be applied to the field of first language acquisition. Finding the answer to questions such as at what stage do
children begin make predictions about the predicates that combine with copulas and what, if any, are the stages in their development of those predictions would further not only the study of copula acquisition in Spanish, but of pragmatic effects in general.

5.4 Final Considerations

The distribution of _ser_ and _estar_ is a classic topic in Hispanic Linguistics and one that has received much attention throughout the years. The main objective of this study is to further our understanding of this distribution, by providing a framework that can accurately describe and formalize adjectival predications based on their scalar properties. In terms of new contributions, the difference in the distribution of copulas between open and closed scale adjectives, and the consequences it has on the online processing of copulas represents a new aspect of the study of _ser_ and _estar_.

ser and estar.
References


Marín, R. (2010). Spanish adjectives within bounds, in P. Cabredo Hofherr & O.


## Appendix A: Target Items for Experiment

<table>
<thead>
<tr>
<th>Item</th>
<th>Ser</th>
<th>Estar with Natural Process</th>
<th>Estar with Artificial Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>Ella está, en este momento, güera.</td>
<td>Ella está, en este momento, colorada.</td>
<td>Ella está, en este momento, pintada roja.</td>
</tr>
<tr>
<td></td>
<td>She is, at this moment, blonde.</td>
<td>She is, at this moment, red.</td>
<td>She is, at this moment, painted red.</td>
</tr>
<tr>
<td>002</td>
<td>El pan está, no lo vas a creer, negro.</td>
<td>El pan está, no lo vas a creer, negro y quemado.</td>
<td>El pan está, you are not going to believe it, black and burned.</td>
</tr>
<tr>
<td></td>
<td>The bread is, you are not going to believe it, black.</td>
<td>The bread is, you are not going to believe it, black and burned.</td>
<td>The bread is, you are not going to believe it, painted black.</td>
</tr>
<tr>
<td>003</td>
<td>El papel es, si te das cuenta, amarillo.</td>
<td>El papel está, si te das cuenta, amarillento.</td>
<td>El papel está, you notice, yellow.</td>
</tr>
<tr>
<td></td>
<td>The paper is, if you notice, yellow.</td>
<td>The paper is, if you notice, yellow.</td>
<td>The paper is, if you notice, painted yellow.</td>
</tr>
<tr>
<td>004</td>
<td>Las manzanas son, si te fías, verdes.</td>
<td>Las manzanas están, si te fías, verdes.</td>
<td>Las manzanas are, if you pay attention, green.</td>
</tr>
<tr>
<td></td>
<td>The apples are, if you pay attention, green.</td>
<td>The apples are, if you pay attention, green.</td>
<td>The apples are, if you pay attention, painted green.</td>
</tr>
<tr>
<td>005</td>
<td>El árbol es, desde luego, rojo.</td>
<td>El árbol está, desde luego, rojo.</td>
<td>El árbol está, since then, red.</td>
</tr>
<tr>
<td></td>
<td>The tree is, doubtless, red.</td>
<td>The tree is, doubtless, red.</td>
<td>The tree is, doubtless, red.</td>
</tr>
<tr>
<td>006</td>
<td>El agua es, increíblemente, verde.</td>
<td>El agua está, increíblemente, verde.</td>
<td>El agua is, incredibly, dyed green.</td>
</tr>
<tr>
<td></td>
<td>The water is, incredibly, green.</td>
<td>The water is, incredibly, green.</td>
<td>The water is, incredibly, dyed green.</td>
</tr>
<tr>
<td>007</td>
<td>El cielo era, cando sali, negro.</td>
<td>El cielo estaba, cando salió, negro de tormenta.</td>
<td>El cielo was, when I left, black from the storm.</td>
</tr>
<tr>
<td></td>
<td>The sky was, when I left, black.</td>
<td>The sky was, when I left, black from the storm.</td>
<td>The sky was, when I left, black from the storm.</td>
</tr>
<tr>
<td>008</td>
<td>El animal es, como puedes ver, azul.</td>
<td>El animal está, como puedes ver, volviéndose azul.</td>
<td>El animal is, as you can see, turning blue.</td>
</tr>
<tr>
<td></td>
<td>The animal is, as you can see, blue.</td>
<td>The animal is, as you can see, turning blue.</td>
<td>The animal is, as you can see, dyed blue.</td>
</tr>
<tr>
<td>009</td>
<td>El río es, no sé cómo, rojo.</td>
<td>El río está, no sé cómo, rojo.</td>
<td>El río is, I don’t know how, red.</td>
</tr>
<tr>
<td></td>
<td>The river is, I don’t know how, red.</td>
<td>The river is, I don’t know how, red.</td>
<td>The river is, I don’t know how, red.</td>
</tr>
<tr>
<td>010</td>
<td>Las flores eran, cómo tú dijiste, negras.</td>
<td>Las flores estaban, cómo tú dijiste, negras por el tiempo.</td>
<td>Las flores were, like you said, black.</td>
</tr>
<tr>
<td></td>
<td>The flowers were, like you said, black.</td>
<td>The flowers were, like you said, black.</td>
<td>The flowers were, like you said, black.</td>
</tr>
<tr>
<td>011</td>
<td>El pelo era, evidentemente, blanco.</td>
<td>El pelo estaba, evidentemente, blanco por las cenizas.</td>
<td>El pelo was, evidently, dyed white.</td>
</tr>
<tr>
<td></td>
<td>The hair was, evidently, white.</td>
<td>The hair was, evidently, white.</td>
<td>The hair was, evidently, dyed white.</td>
</tr>
<tr>
<td>012</td>
<td>Los dientes eran, curiosamente, amarillos.</td>
<td>Los dientes estaban, curiosamente, amarillos.</td>
<td>Los dientes were, curiously, yellow.</td>
</tr>
<tr>
<td></td>
<td>The teeth were, curiously, yellow.</td>
<td>The teeth were, curiously, yellow.</td>
<td>The teeth were, curiously, yellow by the paint.</td>
</tr>
<tr>
<td>013</td>
<td>El puente estaba, curiosamente, rojo.</td>
<td>El puente estaba, curiosamente, rojo por el ácido.</td>
<td>El puente was, curiously, red from the acid.</td>
</tr>
<tr>
<td></td>
<td>The bridge was, curiously, red.</td>
<td>The bridge was, curiously, red.</td>
<td>The bridge was, curiously, red from the acid.</td>
</tr>
<tr>
<td>014</td>
<td>Sus uñas eran, por lo que vi, negras.</td>
<td>Sus uñas estaban, por lo que vi, negras de sucedad.</td>
<td>Sus nails were, by what I saw, black.</td>
</tr>
<tr>
<td></td>
<td>Her nails were, by what I saw, black.</td>
<td>Her nails were, by what I saw, black.</td>
<td>Her nails were, by what I saw, black.</td>
</tr>
<tr>
<td>015</td>
<td>La roca era, definitivamente, verde.</td>
<td>La roca estaba, definitivamente, verde por el moho.</td>
<td>La roca was, definitively, green from the moss.</td>
</tr>
<tr>
<td></td>
<td>The rock was, definitively, green.</td>
<td>The rock was, definitively, green from the moss.</td>
<td>The rock was, definitively, green from the moss.</td>
</tr>
<tr>
<td>016</td>
<td>El coche era, como pude ver, blanco.</td>
<td>El coche estaba, como pude ver, blanco por la nieve.</td>
<td>El coche was, as I could see, white.</td>
</tr>
<tr>
<td></td>
<td>The car was, as I could see, white.</td>
<td>The car was, as I could see, white.</td>
<td>The car was, as I could see, white.</td>
</tr>
<tr>
<td>017</td>
<td>La nariz era, claramente, roja.</td>
<td>La nariz estaba, claramente, roja por el esmalte.</td>
<td>La nose was, clearly, red.</td>
</tr>
<tr>
<td></td>
<td>The nose was, clearly, red.</td>
<td>The nose was, clearly, red.</td>
<td>The nose was, clearly, red.</td>
</tr>
<tr>
<td>018</td>
<td>El ojo era, indubitablemente, negro.</td>
<td>El ojo estaba, indubitablemente, negro por el maquillaje.</td>
<td>El eye was, obviously, black.</td>
</tr>
<tr>
<td></td>
<td>The eye was, obviously, black.</td>
<td>The eye was, obviously, black.</td>
<td>The eye was, obviously, black with make-up.</td>
</tr>
</tbody>
</table>