The Speaker-Oriented Use of *Estar*: Semantic Variation in the [Copula + Adjective] Context in Buenos Aires Spanish

by

Tanya Battersby

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Spanish and Portuguese
University of Toronto

© Copyright by Tanya Battersby 2017
The Speaker-Oriented Use of *Estar*: Semantic Variation in the [Copula + Adjective] Context in Buenos Aires Spanish

Tanya Battersby
Doctor of Philosophy
Graduate Department of Spanish and Portuguese
University of Toronto
2017

Abstract

This study investigates copula variation with adjectival predicates in Buenos Aires Spanish. It examines the conditioning factors of [*ser/estar* + adjective] contexts and tests the hypothesis of the speaker-oriented use of *estar*. The speaker-oriented use of *estar* emphasizes the viewpoint of the speaker (e.g., *La nueva película de Almodóvar está buena*. ‘The new Almodóvar film *ESTAR* good.’). In contrast, *ser* is neutral with respect to the speaker’s attitude towards the proposition (e.g., *La nueva película de Almodóvar es buena*. ‘The new Almodóvar film *SER* good.’). The speaker-oriented hypothesis predicts that the choice of *estar* optionally expresses speaker perspective, which comprises a range of meanings including evidential, mirative and evaluative interpretations. The study tests whether some of the variation between the copulas can be attributed to the encoding of speaker perspective via the selection of *estar*. Data from 45 native speakers of Buenos Aires Spanish was collected, using a contextualized preference task, an elicitation task and a sociolinguistic interview. Speakers were classified according to the two social variables of age and level of education. The findings indicated that the presence of evidence in the discourse context was a significant predictor of *estar* selection. Also, two linguistic variables, the presence of temporal anchoring markers and the presence of speaker-
oriented expressions in the sentence, were highly significant predictors of the use of *estar*. These results are consistent with the prediction that *estar* is selected to express speaker-oriented meaning. Neither of the social variables tested were significant predictors of *estar* selection. The findings suggest stable patterns of use for both temporally-anchored and speaker-oriented meanings of the [*estar* + adjective] context in this variety. In conclusion, the linguistic encoding of speaker perspective is a driver of variation in copula selection with adjectival predicates in Buenos Aires Spanish.
Acknowledgments

This project has only come to fruition by the grace of many people who have helped me along the way. First and foremost, I owe a huge debt of gratitude to my participants, whose words and voices form the backbone of this study. Their warmth, humour, and delight in the art of conversation was immediately apparent in our meetings. I will not forget the hospitality and kindness they showed me in our time together, and the memories of my experiences with them have sustained me throughout many months of transcription, analysis and writing.

I have been fortunate to study with three brilliant, passionate, dedicated linguists, who each inspire in their own way. I extend my heartfelt thanks to Laura Colantoni for her keen observations, unflagging encouragement, and open door. In particular, her support – both moral and practical – during the time of my field work in Buenos Aires was crucial to the success of this project. Cristina Cuervo provided lucid insights into formal syntax and semantics in response to my questions, and our discussions helped illuminate my thinking about key areas of the theoretical aspects of this study. Most of all, I thank Ana Teresa Pérez-Leroux for many years of mentoring. Her detailed feedback and honest opinions have made a tremendous contribution to this thesis. I knew I was on to something whenever I saw her get excited about my research, and her faith in this work compelled me to keep trying to refine my ideas and improve their expression.

In the Department of Spanish and Portuguese, Graduate Administrators Blanca Talesnik and Sara Solís offered every kind of support, help, and encouragement I could ask for. Thanks to Ximena Griscti, José Rozas, Walter Bertolotti, and Silvina Braver for participating in pilot studies for this project and for their help with consulting. Eugenia Pérez Ibáñez and Natalia Elisei shared my recruiting messages with their networks in Buenos Aires, and their efforts were essential to my being able to contact and test many participants in the short period of time that I was there.

During my time in the PhD, I have been lucky to count on the solidarity of my fellow graduate students in the Hispanic Linguistics program, in the Department of Spanish and Portuguese, and in other departments at the University of Toronto. Extra special thanks go to Yadira Álvarez...
López, my counterpart since the beginning in our cohort of two, for all the conversations over the years, throughout which the thread of linguistics was only one of many.

My gratitude to my friends and family for keeping me company on this journey is beyond words. As the Irish phrase goes, *Mo sheasamh ort lâ na choise tinne*, “You are the place where I stand on the day when my feet are sore.”¹ Finally, I am infinitely thankful to my Mum and Dad, for encouraging my curiosity in the first place. There is no way I could have done this without them.

# Table of Contents

Table of Contents ........................................................................................................ iv
List of Tables .................................................................................................................. viii
List of Figures ................................................................................................................. xi
List of Appendices .......................................................................................................... xii

Chapter 1 Introduction .................................................................................................. 1

Chapter 2 Syntax and semantics of *ser* and *estar* with adjectives .......................... 4

  2.0 Introduction ........................................................................................................... 4
  2.1 The distribution of *ser* and *estar* ................................................................... 4
  2.2 Copula theory ..................................................................................................... 11
  2.3 Accounts of *ser* and *estar* .......................................................................... 12
    2.3.1 Aspectual approaches: Aspect as a binary lexical feature ....................... 12
    2.3.2 The predicate-type distinction ................................................................. 14
    2.3.3 Pragmatics ............................................................................................... 20
    2.3.4 An asymmetrical aspectual analysis ....................................................... 24
    2.3.5 Prepositional explanations .................................................................... 27
  2.4 Properties of the adjective .................................................................................. 30
    2.4.1 Semantic classification of adjectives ....................................................... 30
    2.4.2 The stative/eventive contrast ................................................................ 32
    2.4.3 Gradability ............................................................................................. 35
    2.4.4 Summary: Properties of the adjective ................................................... 43
  2.5 The evidential use of *estar* .............................................................................. 44
    2.5.1 Evidential coercion with *estar* ............................................................. 45
    2.5.2 The discovery interpretation of *estar* ................................................... 46
    2.5.3 Implied comparisons with *estar* ........................................................... 48
    2.5.4 Summary: The evidential use of *estar* .................................................. 51
  2.6 Evidentiality ........................................................................................................ 52
    2.6.1 Definitions of evidentiality ..................................................................... 52
    2.6.2 Evidential morphosyntax ..................................................................... 54
    2.6.3 The internal structure of evidential systems ........................................ 56
    2.6.4 Evidentiality and related meanings ....................................................... 60
    2.6.5 Semantics of evidentials ...................................................................... 61
    2.6.6 Summary: Evidentiality ....................................................................... 62

Chapter 3 Variation and change in the use of *ser* and *estar* with adjectives ............ 64

  3.0 Introduction ........................................................................................................... 64
  3.1 Diachronic studies ............................................................................................... 64
  3.2 Recent studies: Methodologies .......................................................................... 66
    3.2.1 Populations ............................................................................................. 66
    3.2.2 Quantitative approaches ...................................................................... 67
  3.3 Results of previous studies: Linguistic variables ............................................... 74
    3.3.1 Linguistic variables: Item-level variables .............................................. 74
    3.3.2 Linguistic variables: Sentence-level variables ....................................... 79
    3.3.3 Linguistic variables: Discourse-level variables ..................................... 83
    3.3.4 Summary of linguistic variables ............................................................. 87
  3.4 Results of previous studies: Extralinguistic variables ........................................ 89
List of Tables

Table 2.1. Summary of the distribution of *ser* and *estar*

Table 2.2. Quantized structures (Husband, 2012, p. 151)

Table 2.3. Types of evidence (Willett, 1988, p. 57)

Table 2.4. Semantic parameters in evidentiality systems (Aikhenvald, 2004, p. 65)

Table 2.5. Three dimensions of evidential meaning (Matthewson, 2011, p. 4)

Table 3.1. Previous studies: Populations and data collection methods

Table 3.2. Parameters for copula selection (Silva-Corvalán, 1986, p. 590)

Table 3.3. Summary of linguistic variables: Categories and findings

Table 3.4. Summary of extralinguistic factors: Variables and findings

Table 4.1. Design: Contextualized Preference Task

Table 4.2. Adjectives by Condition: Contextualized Preference Task

Table 4.3. Distractors and Practice Items: Contextualized Preference Task

Table 4.4. Overview of participant data by age and level of education

Table 4.5. Controlled Elicitation Task: Visual Stimuli: Images versus Recalled Images

Table 4.6. Controlled Elicitation Task: Auditory Stimuli: Sounds versus Recalled Sounds

Table 4.7. Controlled Elicitation Task: Olfactory Stimuli: Scents versus Recalled Scents

Table 4.8. Controlled Elicitation Task: Tactile Stimuli: Objects versus Recalled Textures

Table 4.9. Controlled Elicitation Task: Gustatory Stimuli: Foods versus Recalled Tastes

Table 5.1. Frequency of selection of *ser* and *estar*/both according to presence of evidence in the context

Table 5.2. Frequency of selection of *ser* and *estar*/both by evidence type in evidence-present versus evidence-neutral contexts

Table 5.3. Frequency of selection of *ser* and *estar*/both in evidence-present versus evidence-neutral contexts according to speaker age
Table 5.4. Frequency of selection of *ser* and *estar* both in evidence-present versus evidence-neutral contexts according to speaker level of education

Table 5.5. Mixed-effects logistic regression analysis of the contribution of external and internal factors to the probability of *estar* selection in the contextualized preference task with Speaker and Item as random effects

Table 5.6. Controlled Elicitation Task: Summary of adjective coding

Table 5.7. Controlled Elicitation Task: Overall distribution of *ser* and *estar* with qualificational adjectives

Table 5.8. Controlled Elicitation Task: Distribution of *ser* and *estar* with qualificational adjectives according to evidence type

Table 5.9. Controlled Elicitation Task: Distribution of *ser* and *estar* with qualificational adjectives according to evidence and type of sensory stimulus

Table 5.10. Controlled Elicitation Task: Distribution of *ser* and *estar* with qualificational adjectives according to age and evidence type

Table 5.11. Controlled Elicitation Task: Distribution of *ser* and *estar* with qualificational adjectives according to level of education and evidence type

Table 5.12. Mixed-effects logistic regression analysis of the contribution of external and internal factors to the probability of *estar* selection with qualificational adjectives in Buenos Aires Spanish with Speaker and Adjective as random effects

Table 5.13. Controlled Elicitation Task: Distribution of adjectives with *ser* and *estar* by evidence type for adjectives with >1 token of *estar* selection in direct evidence contexts

Table 5.14. Interview Task: Overall distribution of *ser* and *estar* with qualificational adjectives

Table 5.15. Overall distribution of *ser* and *estar* with qualificational adjectives in interview data and elicitation data combined

Table 5.16. Summary of coding of speaker-oriented expressions

Table 5.17. Combined interview and elicitation data: Distribution of *ser* and *estar* with qualificational adjectives according to presence of temporal anchoring markers in the utterance

Table 5.18. Combined interview and elicitation data: Distribution of *ser* and *estar* with qualificational adjectives according to presence of speaker-oriented expressions in the utterance

Table 5.19. Combined interview and elicitation data: Distribution of *ser* and *estar* with qualificational adjectives according to presence of speaker-oriented expressions and temporal anchoring markers in the utterance
Table 5.20. Mixed-effects logistic regression analysis of the contribution of external and internal factors to the probability of *estar* selection with qualificational adjectives in Buenos Aires Spanish with Speaker and Adjective as random effects
List of Figures

Figure 5.1. Combined interview and elicitation data: Use of *ser* and *estar* with qualificational adjectives by age according to presence [+/-] of speaker-oriented expressions in the utterance

Figure 5.2. Combined interview and elicitation data: Use of *ser* and *estar* with qualificational adjectives by age according to presence [+/-] of temporal anchoring markers in the utterance

Figure 5.3. Combined interview and elicitation data: Use of *ser* and *estar* with qualificational adjectives by level of education according to presence [+/-] of speaker-oriented expressions in the utterance

Figure 5.4. Combined interview and elicitation data: Use of *ser* and *estar* with qualificational adjectives by level of education according to presence [+/-] of temporal anchoring markers in the utterance
List of Appendices

Appendix A: Stimuli: Contextualized Preference Task
Appendix B: Language and Personal Background Questionnaire
Appendix C: Participant Data
Appendix D: Stimuli: Controlled Elicitation Task
Appendix E: Script: Guided Interview Task
Chapter 1
Introduction

This study examines variation between the copula verbs *ser* and *estar* with adjectival predicates, as in (1).

(1) a. *Javier es/está nervioso.*
    ‘Javier is SER/ESTAR nervous.’

b. *El día es/está bonito.*
    ‘The day is SER/ESTAR beautiful.’

The [copula + adjective] context is an enduring topic of interest in both formal theories of syntax and semantics and in studies of variation and change in Spanish. The question of variation and change in the use of *ser* and *estar* with adjectives engages multiple aspects of language, from syntax to semantics to pragmatics.

Patterns of variation and change with *ser* and *estar* present a unique opportunity to gain insight into the grammar. When a meaning is lost, or a new one acquired, there are potential consequences in different domains of grammar. For example, syntactic proposals of the Spanish copula contrast disagree as to whether there is a structural parallel between *ser* and *estar*, and evidence from empirical studies of the copulas can contribute to this debate. How speakers use [copula + adjective] sentences in spontaneous speech illustrates differences in possible interpretations. The properties of the adjectives and the subjects in these sentences reveal the way meaning is constructed through semantic composition. Restrictions on the types of adjectives or subjects that combine with *ser* and *estar* may be observed to vary or change, and in such cases, the data presented in studies of variation and change provide a basis for the refinement of theories about the different components of the sentence and how they interact. Finally, the existence of simultaneous competing interpretations of certain *ser/estar* clauses suggests that the discourse context plays a role in the determination of meaning. Variation in the
[copula + adjective] context draws attention to pragmatics as a source of semantic change through reanalysis. In sum, this topic opens a window onto issues such as the nature and properties of copulas and adjectives, the mapping between predicates and their subjects, the quantization of states, the role of pragmatic context in the construal of meaning, and the questions of how and why structures and meanings vary and change.

The goal of this study is to make a novel contribution to these topics through the investigation of the use of \([ser/estar + \text{adjective}]\) clauses in Buenos Aires Spanish. As far as I am aware, there are no previous studies of variation and change in copula selection that focus on this dialect. This study uses both experimental and sociolinguistic interview data to add to the overall descriptive picture of this phenomenon across the Spanish-speaking world. In addition, I examine a meaning that has been discussed in the theoretical literature but which has not been approached from an empirical perspective until now: the evidential use of \(estar\). Previous sociolinguistic research on variation and change in the \([ser/estar + \text{adjective}]\) context has investigated the semantic extension of \(estar\), but has largely overlooked the role of evidential factors affecting copula choice with adjectives. I argue that some of the variation between the copulas can be attributed to the encoding of the speaker’s perspective via the selection of \(estar\). The evidential use of \(estar\), as in (2), emphasizes the viewpoint of the speaker. In contrast, \(ser\) is neutral with respect to the speaker’s attitude towards the proposition, as in the contexts in (3).

\[(2)\]
\begin{enumerate}
  \item \(Este vino está riquísimo.\)
  
  ‘This wine \text{i\_ESTAR delicious\_INT.}.’
  
  \item \(¿Viste la nueva película de Almodóvar? Está buena.\)
  
  ‘Did you see the new Almodovar film? \text{i\_ESTAR good.’}
  
  \item \(Están muy lindos los cachorros del vecino.\)
  
  ‘The neighbor’s puppies \text{a\_ESTAR cute.’}
\end{enumerate}

\[(3)\]
\begin{enumerate}
  \item \(Este vino es riquísimo.\)
  
  ‘This wine \text{i\_SER delicious\_INT.’}
\end{enumerate}
b. ¿Viste la nueva película de Almodóvar? Es buena.

‘Did you see the new Almodovar film? It is \textit{SER} good.’

c. Son muy lindos los cachorros del vecino.

‘The neighbor’s puppies are \textit{SER} cute.’

I identify the evidential use of \textit{estar} as a driver of variation and change that should be distinguished from the contexts of semantic bleaching of the copula contrast described elsewhere as the innovative use of \textit{estar}. I also incorporate generalizations from the theoretical and typological literature on evidentiality into the investigation of the evidential use of \textit{estar}, and inquire into what kinds of evidential meaning are encoded by this form.

The following is an outline of the study. In Chapter 2, I review and critically evaluate the theoretical and empirical background literature on copulas, adjectives, and \textit{ser} and \textit{estar} predicates. I also consider the contribution of the literature on evidentiality to an explanation of the evidential use of \textit{estar}. I examine the findings of previous studies of variation and change in the [copula + adjective] context in Spanish in Chapter 3. Next, in Chapter 4, I develop a hypothesis about the evidential use of \textit{estar} based on formal theories of evidentiality across languages, and then present the methodology of the study. Chapter 5 reports the results of the three experimental tasks: a contextualized preference task, a controlled elicitation task and an interview task. I also propose revisions to my original hypothesis and discuss my revised characterization, the speaker-oriented use of \textit{estar}. Chapter 6 concludes and suggests avenues for further investigation.
Chapter 2
Syntax and semantics of *ser* and *estar* with adjectives

2.0 Introduction

This goal of this chapter is to examine the alternation between *ser* and *estar* with adjectival predicates. I consider the structure and meaning of both copula and adjective, and review the main proposals that have attempted to explain the puzzle of *ser* and *estar*. I also evaluate previous approaches to the evidential use of *estar*, and finally, summarize relevant aspects of the background literature on evidentiality.

2.1 The distribution of *ser* and *estar*

*Ser* and *estar* have a mixed distribution, with some categorical contexts and some variable contexts. Variable contexts include sentences with Adjective Phrase (AP) complements and locative sentences with subjects that are immovable objects. Adjectival predicates can be broadly divided into three categories: adjectives that only combine with *ser*, those that only combine with *estar*, and adjectives that combine with either *ser* or *estar* (Luján, 1981). The majority of adjectives belong to this third category, and Vaño-Cerdá (1982) estimates that 80% of all adjectives in Spanish may appear with either copula. The different readings of *ser* and *estar* predicates are numerous, and I present the best-known, and perhaps least controversial, interpretations as a point of departure.

First, there are types of adjectives that occur exclusively with *ser*, such as denominal adjectives (4). These sentences express an atemporal predication of a property of an individual/entity.

(4)  *Javier es/*está argentino.*
'Javier is*SER/*ESTAR Argentine.'

Then, there are other adjectives that only appear with *estar* (5). These sentences produce a stative reading in which the predicate applied to an individual/entity is subject to spatiotemporal limitation.

(5)  

\[ \text{Javier \*es/está descalzo.} \]

‘Javier is*SER/*ESTAR barefoot.’

Finally, there are adjectives with variable behaviour. Some of these adjectives, like *vivo* (‘sharp’/‘alive’) or *listo* (‘clever’/‘ready’), are cases of polysemy, where the meanings of the adjectives that combine with one copula or the other can be attributed to distinct, semantically-related lexical entries (6, 7).

(6)  

\[ \text{Javier es listo.} \]

‘Javier is*SER clever.’

(7)  

\[ \text{Javier está listo.} \]

‘Javier is*ESTAR ready.’

However, many adjectives can appear with either *ser* or *estar*, with a difference in meaning that arises from its combination with one verb or the other (8, 9). These sentences are differentiated from cases of polysemous adjectives; it is the verbs that produce the contrast between property and state meanings in (8, 9), and the adjective is assumed to have a single lexical entry.

(8)  

\[ \text{Javier es nervioso.} \]

‘Javier is*SER nervous.’

(9)  

\[ \text{Javier está nervioso.} \]

‘Javier is*ESTAR nervous.’
The sentence with *ser* in (8) generates a property interpretation. Nervousness is a stable, inherent property of Javier in (8). In other words, Javier is generally a nervous person. The sentence with *estar* in (9) produces a stative reading. That is, in (9), Javier’s current condition is one of nervousness.

The verbs *ser* and *estar* mark a contrast between verbal and adjectival passive sentences in Spanish. When *ser* is followed by a past participle, an eventive reading is obtained (10). *Estar* with a past participle gives rise to a stative reading (11). The sentence in (11) describes a resultant state.

(10)  *La casa fue destruida (por la tormenta).*

‘The house was*SER* destroyed (by the storm).’

(11)  *La casa estaba destruida (*por la tormenta).*

‘The house was*ESTAR* destroyed (*by the storm).’

To sum up, both *ser* and *estar* can combine with APs, with distinct interpretations. Additional readings of adjectival predicates with *ser* and *estar* are available, and multiple factors are involved in the interpretations of these sentences.

The classic intuition is that the difference between sentences where the distribution of *ser* and *estar* overlaps relates to the spatio-temporal effects of the predicate. A traditional explanation is that the semantic contrast can be attributed to the expression of a temporary state (*estar*) versus a permanent property (*ser*) (Bello, 1847). Thus, in (4), repeated here, *ser* is selected to convey that Javier’s being Argentine is a permanent characteristic, whereas the choice of *estar* in (5), repeated here, suggests that Javier’s being barefoot is a temporary condition.

(4)  *Javier es/*está argentino.*

‘Javier is*SER/*ESTAR Argentine.’

(5)  *Javier *es/está descalzo.*

‘Javier is*SER/ESTAR barefoot.’
This is comparable to another traditional description, that of inherent quality (*ser*) versus current condition (*estar*) (Gili y Gaya, 1972). In a similar vein, the analysis of *estar* as an expression of a change of state (Fernández-Leborans, 1995) involves a change-of-state meaning ascribed to the predicates in *estar* sentences, whereas the states in which the subjects in *ser* sentences find themselves are unrestricted by time. In (12), Javier’s tallness applies without regard to time or place. In (13), however, the use of *estar* indicates a change of state that applies to Javier’s height.

(12)  *Javier es alto.*

‘Javier is *SER* tall.’

(13)  *Javier está alto.*

‘Javier is *ESTAR* tall.’

Another type of [copula + adjective] clause is the evidential use of *estar*. The sentences in (2), repeated here, are examples of this context.

(2)  a.  *Este vino está riquísimo.*

‘This wine is *ESTAR* delicious*INT*.’

b.  *¿Viste la nueva película de Almodóvar? Está buena.*

‘Did you see the new Almodovar film? It is *ESTAR* good.’

c.  *Están muy lindos los cachorros del vecino.*

‘The neighbor’s puppies are *ESTAR* cute.’

The basic insight is that *estar* predication in these cases is related to the viewpoint of the speaker, rather than any of the other meanings that the choice of *estar* over *ser* could generate.

The evidential use of *estar* stands as an exception to traditional explanations. The meaning conveyed in the construction has many of the qualities of a prototypical *ser* predication. The evidential use of *estar* attributes an atemporal property to the subject – a context where, prescriptively, we would expect to see *ser*, as in (14).

    ‘The wine is SER delicious.’

  b.  *La película es buena.*

    ‘The film is SER good.’

  c.  *Los cachorros son lindos.*

    ‘The puppies are SER cute.’

Nor are evidential uses of *estar* limited to a stage that is associated with a particular location or time (15b,c). That is, it does not express a state, unlike many *estar* predications (15a).

(15)  a.  *El parque está lleno el fin de semana.*

    ‘The park is ESTAR full on the weekend.’

  b.  #*El vino está rico el fin de semana.*

    #‘This wine is ESTAR delicious on the weekend.’

  c.  #*La película está buena el fin de semana.*

    #‘The film is ESTAR good on the weekend.’

Nor is the predication linked to a change in its recent past or near future (16b), as some predicates are with *estar* (16a).

(16)  a.  *Javier está triste hoy, pero no lo estaba ayer/no lo estará mañana.*

    ‘Javier is ESTAR sad today, but he wasn’t ESTAR yesterday/he won’t be ESTAR tomorrow.’

  b.  #*La nueva película de Almodóvar está buena hoy, pero no lo estaba ayer/no lo estará mañana.*
"The new Almodovar film is ESTAR good today, but it wasn’t ESTAR yesterday/it won’t be ESTAR tomorrow."

In addition, [estar + adjective] clauses with resultative (17) and change-of-state readings (18) can give rise to ambiguity with evidential meanings. In (17), the presence of the perfective adjective oscuro ‘dark’ suggests a resultative meaning, but discourse-contextual factors (such as the speaker looking up at the sky at the time of speech) could generate an accompanying evidential meaning. Similarly, in (18), the speaker’s observation of Javier’s height might encode surprise (i.e., mirative meaning, a type of meaning related to evidential meaning), in addition to the marking of a change of state expressed by the choice of estar over ser.

(17) El cielo está oscuro.

‘The sky is ESTAR dark.’

(18) Javier está alto.

‘Javier is ESTAR tall.’

The evidential use of estar requires an explanation, but few approaches to the copula distinction include it in their accounts.

More generally, the key question in the literature on ser and estar is concerned with whether the copula or the adjectival predicate is responsible for the differing interpretations of clauses with these verbs. On one hand, sentences where an adjective is compatible with both copulas (e.g., 8-9, and 12-13) would suggest that the copula is the crucial element in generating the semantic contrast. On the other hand, if we consider that some predicates may only appear with one copula (e.g., 4-5), we might be inclined to attribute some kind of grammatical role to the adjective. I examine both copula and adjective in detail and consider their respective contributions to the [copula + adjective] structure in the sections that follow. To conclude this section, I present an overview of the main syntactic constraints and interpretational distinctions between the copulas in Table 2.1. The table is adapted from the discussion in Fernández-Leborans (1999). Examples are my own.
Table 2.1. Summary of the distribution of *ser* and *estar*

<table>
<thead>
<tr>
<th>Post-copular constituent</th>
<th>With <em>ser</em></th>
<th>With <em>estar</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>AP (neither type 1 or 2 exclusively)</td>
<td>√ Javier es nervioso.</td>
<td>√ Javier está nervioso.</td>
</tr>
<tr>
<td>AP (evidential meaning)</td>
<td>*El vino es rico.</td>
<td>√El vino está rico.</td>
</tr>
<tr>
<td>Gerund (progressive tense)</td>
<td>* Flavia es haciendo un esfuerzo grande.</td>
<td>√ Flavia está haciendo un esfuerzo grande.</td>
</tr>
<tr>
<td></td>
<td>* Mariano era laburando todo el día.</td>
<td>√ Mariano estaba laburando todo el día.</td>
</tr>
<tr>
<td>Verbal passive</td>
<td>√ La casa fue destruída (por la tormenta).</td>
<td>* La casa estaba destruída (por la tormenta).</td>
</tr>
<tr>
<td>Adjectival passive</td>
<td>* La casa fue destruída.</td>
<td>√ La casa estaba destruída.</td>
</tr>
<tr>
<td>Locative PP (movable object)</td>
<td>* La pelota es en el jardín.</td>
<td>√ La pelota está en el jardín.</td>
</tr>
<tr>
<td>Locative PP (immovable object)</td>
<td>√ Mi casa es por allí.</td>
<td>√Mi casa está por allí.</td>
</tr>
<tr>
<td></td>
<td>√ La frazada es de lana de alpaca.</td>
<td>* La frazada está de lana de alpaca.</td>
</tr>
<tr>
<td>Possessive PP</td>
<td>√ La frazada es de Paula.</td>
<td>* La frazada está de Paula.</td>
</tr>
</tbody>
</table>
2.2 Copula theory

The goal of many theoretical proposals is to capture the interpretations of [copula + adjective] clauses in both categorical and variable copula contexts, while at the same time maintaining a unified structural account of the elements involved. In order to understand the contribution of each element to the sentences at issue here, I will start by considering the role of the copula.

The diversity of copula forms across languages is by now well-documented by typological studies (e.g. Pustet, 2003; Stassen, 1997). However, these studies were preceded by a lively historical debate over the functions of copulas. Historically, three main perspectives on the nature and function of the copula can be identified: the copula as a semantically null carrier of inflection, the copula as a predicative linking mechanism, and the copula as the expression of an identity function between subject and complement (Moro, 1997). These views continue to surface in different guises in recent theories of the syntactic structure and semantic composition of copula clauses. Underlying the theory of copula clauses are fundamental questions about phrase structure and the interrelationships between syntax, semantics and pragmatics.

Higgins’s (1979) seminal proposal of a taxonomy classified copular clauses according to four functions: predicational, specificational, identificational and identity. In subsequent work, the functions of copular clauses have been analyzed from a variety of vantage points. Three broad perspectives can be described. The first point of view (e.g. Adger & Ramchand, 2003; Bowers, 1993; den Dikken, 2006; Moro, 1997) argues for a single underlying structure for copular clauses. The different clause types enumerated by Higgins illustrate the various surface realizations of this underlying structure. A second point of view regards the clause types proposed by Higgins as different readings that may obtain of copular clauses, depending on the meanings of the elements and their semantic composition (e.g. Partee, 1986, 1987; Rothstein, 1995, 1999, 2001). The third perspective considers the interaction between formal elements and pragmatic factors in producing the different types of copular functions (e.g. Heller, 2005; Mikkelsen, 2005). The sentences types given in Higgins’s taxonomy represent specific syntax-semantics configurations that vary according to their information structures. This is perhaps the
most faithful approach to the original intention behind Higgins’s taxonomy, as Higgins himself was truly conscious of and concerned with the pragmatics of copular clauses:

The factor which seems to be important in copular sentences is more often a distinction between what is known and is familiar and what is not known or is unfamiliar. Because of this, the copular sentence plays an essential role in the communication of new information about known things… (1973, p. 193)

The flexibility of the copula to enter into a wide range of configurations also means that the discourse context perhaps plays a more important role in the construction of meaning in these clauses than in structures with less flexible elements, as is the case for many of the contexts in which ser and estar appear.

2.3 Accounts of ser and estar

2.3.1 Aspectual approaches: Aspect as a binary lexical feature

In her account of the Spanish copulas, Luján (1981) considers the aspectual dimension to the semantic contrast expressed by ser and estar. She argues that the distinction between the copulas can be attributed to their aspectual nature, where ser is the [-PERFECTIVE] copula and estar is the [+PERFECTIVE] copula. The adjectival predicates that combine with the copulas are also specified for aspect. Adjectives that combine with ser are [-PERFECTIVE], adjectives that combine with estar are [+PERFECTIVE], and adjectives that combine with either copula are [-PERFECTIVE]. For Luján, a ser predicate is made up of a series of intervals, or estar predicates. However, the assumption cannot be reversed: ser predication implies estar predication, but estar predication does not imply ser predication. Luján maintains that this partial synonymy of the copulas is captured by the fact that both can be stative, whereas the meaning contrasts that obtain are attributed to the [+/-PERFECTIVE] feature of the adjectival predicate. One problem with Luján’s account is its view of perfectivity as a lexical, rather than functional, feature. Her work
has also been criticized for its conflation of the aspectual layers where perfectivity (the level of grammatical aspect, outer aspect or viewpoint aspect) and telicity (referred to as Aktionsart, inner aspect or situation aspect) are determined. Although subsequent analyses depart from Luján’s conceptualization of aspect, her proposal lays an important foundation for further development of aspectual theories about the copula distinction.

Roby (2009) also explains the copula contrast in terms of aspectual features, and proposes that *ser* and *estar* express the same binary [+/-PERFECTIVE] distinction as the Spanish preterite and imperfect tenses. He follows in the tradition of Luján in subscribing to a view of aspect as a lexical feature, while at the same time adopting a configurational approach to aspect. With respect to the latter, Roby applies a framework in which there are three levels of aspect. The lexicon represents the first level, where aspectual information is encoded in the form of [+/-PERFECTIVE] features. *Ser* is [-PERFECTIVE] and *estar* is [+PERFECTIVE]. The second level, inner aspect, concerns the readings that result from the interaction between the lexical items in a sentence. It is here that (a)telicity is determined according to the features of the verb and its arguments. Finally, the third level of aspect exists at the level of phrase structure and is referred to as outer aspect. Grammatical [+/-PERFECTIVE] features appear in the domain of INFL. The preterite tense expresses the [+PERFECTIVE] feature, and the imperfect tense expresses the [-PERFECTIVE] feature. The inflectional morphology that this domain contributes is independent of the aspectual information expressed at the level of inner aspect, as well as the aspectual features that come from the lexicon. Roby argues that the aspectual information expressed by INFL has scope over the entire sentence. In other words, it determines the overall [+/-PERFECTIVE] interpretation of the sentence.

A problem for Roby’s account is that there are no data to suggest that the functional [+/-PERFECTIVE] feature of the verb phrase has scope over the lexical [+/-PERFECTIVE] feature of the copula, if such a thing exists. Both *ser* and *estar* can combine with either tense, as in (19).

(19) a.   *El día fue bonito.*

‘The day was* ser,PRET beautiful.’

b.   *El día era bonito.*
‘The day was **SER.IMP** beautiful.’

c. *El día estuvo bonito.*

‘The day was **ESTAR.PRET** beautiful.’

d. *El día estaba bonito.*

‘The day was **ESTAR.IMP** beautiful.’

The meaning expressed by these sentences is the result of semantic composition among the lexical items (in particular, the quantization of the predicate that results from the specific [copula + adjective] combination), and the grammatical aspect contributed by the preterite or imperfect tense. What arises out of the interaction among these elements is a partial neutralization of the meaning distinction between the copulas, rather than the grammatical aspect taking scope over other elements of meaning.

### 2.3.2 The predicate-type distinction

Another approach that is conceptually linked to traditional notions of a temporary/permanent contrast in the semantics is the distinction between predicate types, based on Carlson’s (1977, 1980) proposal of a semantic classification for situation types in terms of Individual Level (IL) and Stage Level (SL) predicates. IL predicates denote properties of individuals, whereas SL predicates denote properties of stages of individuals. A stage is defined as a spatially or temporally bounded manifestation of an individual; an individual is what ties a series of stages together as one thing (Carlson, 1980, p. 68). According to Carlson, IL predicates (20a) have generic readings, whereas SL predicates (20b) may give rise to generic or existential readings.

(20) a. Doctors are intelligent.

    b. Doctors are available.

(1980, p. 74)
These generalizations about predicates that express permanent or essential qualities as IL, and predicates that express transitory or accidental qualities as SL, laid the groundwork for an application of the predicate-type analysis to the copula distinction in Spanish.

Numerous approaches examine the possibility that *ser* and *estar*, respectively, could be an expression of the IL/SL predicate distinction, with varying implementations of the theory (Arche, 2006; Diesing, 1992; Fábregas, 2012; Fernández-Leborans, 1999; Kratzer 1995). The contrasting interpretations of these two types of predicates are explained in terms of differences in their argument structure in Diesing (1992) and Kratzer (1989, 1995). This syntactic account of the differences in IL versus SL predication contains two key components. The first is Kratzer’s proposal that SL predicates possess an eventive or spatiotemporal argument which IL predicates lack. The second is Diesing’s split-VP hypothesis, in which the subjects of IL and SL predicates are mapped onto different syntactic positions. Subjects of IL predicates are generated in IP, and thus may only receive generic readings (21). SL predicates have an extra argument position for events or spatiotemporal locations that IL predicates lack, and which allows subjects to be generated in VP. This means that the subjects of SL predicates can either raise to IP to produce a generic reading or stay in the VP for an existential reading (22).

(21) IL Predicate structure

```
  IP
   /\     \\
  DP_i I'  \\
     /\     \\
    I° VP
        \
       /\     \\
      PRO_i V'  \\
         /\     \\
        V° XP
```
Several proposals explore the mapping between the *ser/estar* contrast and the IL/SL distinction. Fernández-Leborans (1999) brings together the two main theoretical approaches to the *ser/estar* puzzle – the aspectual analysis and the predicate-type distinction. In her account, *ser* is only compatible with IL predicates and *estar* appears exclusively with SL predicates. This distribution is an expression of the lexical-semantic aspectual properties of the verbs. Because *estar* possesses internal temporal-aspectual structure, it is compatible with SL predicates, which have spatio-temporal limitations. *Ser*, on the other hand, lacks internal temporal structure and is aspectually “inert”, which explains its affinity for IL predicates. IL predicates express properties without any temporal-aspectual limitation on the subject. Fernández-Leborans stops short of absolute generalizations, however, in claiming that most (but not all) adjectives that combine with *estar* are perfective, owing to the implication of a termination point in their semantics. With respect to the adjectives that are classified as IL predicates and are not perfective, Fernández-Leborans argues that they behave as SL predicates when they combine with *estar*. The requisite element for an adjective to be able to combine with *estar* is that it must denote a property that has the capacity to change over time.

Arche (2006) also sees the validity in both the IL/SL predicate contrast and an aspectual approach in terms of explaining the behaviour of *ser* and *estar*. She looks to the syntax to determine how the mechanics of the aspectual differences might function, while at the same time assuming that the predicate-type distinction corresponds directly to the semantics of *ser* and *estar*. Arche argues that the copula contrast, and more generally the contrast between IL and SL
predications, can be attributed to the structural layering of tense, inner aspect and outer aspect. Each layer is responsible for a different kind of temporal contribution to the overall structure and interpretation. Tense establishes an ordering between Topic Time and Reference Time\(^2\). Inner aspect refers to the mereological properties of predicates, and distinguishes between homogeneous and heterogeneous predicates. Outer aspect is a “complex projection containing a quantifier” that expresses the “different number of occasions eventualities occur and different orderings between the Eventuality Time and the Topic Time intervals” (Arche, 2006, p. 242).

Arche maintains a strong commitment to the relevance of the IL/SL distinction in capturing the differences between \textit{ser} and \textit{estar}, and takes the position that all adjectives are IL “by default”. More precisely, she suggests that the IL/SL dichotomy does not correspond directly to the notions of permanence/temporariness. In her view, predicate type (IL or SL) is established at the level of inner aspect, and aspectual operators (progressive or perfective) correspond to outer aspect.

In addition, Arche points out that the \textit{ser/estar} contrast originates in the lexicon. The distinct lexical properties of each verb are responsible for the interpretational differences produced. Arche describes these differences as the general intuition of native speakers that \textit{estar} clauses link a predication to a specific situation, whereas \textit{ser} serves as a means to assign a property to an individual. The lexical meaning of \textit{estar} is the source of the association of the predicate with a specific situation, whereas \textit{ser} is the lexical exponent of predication. It is semantically “vacuous” and behaves like a copula in the Aristotelian sense.

The IL/SL explanation of the patterning of \textit{ser} and \textit{estar} with adjectives has been criticized extensively (e.g. Maienborn, 2005; Schmitt, 2005), suggesting that the assumption of a direct correspondence between the IL/SL distinction and the \textit{ser/estar} contrast is not viable. First, although the subject of an IL predicate with \textit{ser} can only receive a generic interpretation (23a), an SL predicate with \textit{estar} can be ambiguous between generic and existential interpretations (23b), but only with an indefinite determiner.

\(^2\) In Arche’s proposal, the terms Topic Time (TT), Reference Time (RT) and Eventuality Time (ET) are informed by the work of Reichenbach (1947), Klein (1994), Zagona (1990) and Stowell (1993, 1996). In brief, the temporal interpretation of a sentence can be conceived of in terms of the relationships among three intervals: the time of the event (ET), a time of reference (RT) and “the interval the speaker talks about” (TT) (Arche, 2006, pp. 151-152).
(23)  a.  *Un bombero es altruista.  (generic)

      ‘A firefighter isSER altruistic.’

b.  Un bombero está disponible.  (generic and existential)

      ‘A firefighter isESTAR available.’

c.  El bombero está disponible.  (existential)

      ‘The firefighter isESTAR available.’

This ambiguity is restricted to indefinite subjects, whereas definite subjects produce only an existential reading with estar (23c), suggesting that the source of the ambiguity can be traced to the determiner rather than the copula, as proposed by Sánchez and Camacho (1992).

Additional counterevidence to the IL/SL characterization of ser and estar comes from Maienborn (2005). She argues that neither IL nor SL predicates have an eventuality argument, and presents data demonstrating that neither ser nor estar predications are acceptable in three standard diagnostic tests for eventuality expressions. These tests include the ability to combine with locative modifiers (24) and manner adverbials (25), and to appear as infinitival complements of perception verbs (26).


      ‘The toy isSER yellow under the tree.’

b.  *La camisa está mojada sobre la silla.

      ‘The shirt isESTAR wet on the chair.’

      (Maienborn, 2005, pp. 162-163)

(25)  *Las manzanas eran/estaban dulces sabrosamente.

      ‘The apples wereSER/ wereESTAR sweet deliciously.’
(Maienborn, 2005, p. 165)

(26) ??Yo vi a Carol ser/estar guapa.

‘I saw Carol be\textit{ser}/be\textit{estar} beautiful.’

(Maienborn, 2005, p. 166)

Maienborn concludes that neither \textit{ser} nor \textit{estar} introduces a Davidsonian eventuality argument, contrary to the predictions of the IL/SL predicate-type distinction. These observations suggest that the IL/SL hypothesis for \textit{ser} and \textit{estar} is simply not supported by the data.

Given the extensive counterevidence to a theory of \textit{ser} and \textit{estar} as lexical exponents of the IL/SL distinction, is there anything to be gained by continuing to pursue this analysis? This is the question behind Fábregas’s (2012) work, in which he argues that the predicate types are not syntactic primitives, but rather represent a cluster of lexical, structural and conceptual properties. He makes reference to two basic criteria that underlie the distinction as originally conceived by Milsark (1974). The first is whether or not the quality is intrinsic to the identity of the individual, and the second is the presence or absence of temporal boundedness. A case like the sentence in (27) demonstrates that even an archetypal IL predicate expressing nationality (which is expected to be intrinsic to the individual as well as atemporal) can receive an SL reading in the right context.

(27) \textit{Juan se fue a Alemania español, pero volvió alemán.}

‘Juan went to Germany Spanish, but he came back German.’

(Fábregas, 2012, p. 7)

Fábregas goes on to ask whether the inverse scenario might also be possible. Could a stage-level adjective be acceptable an individual-level context (28)? He reasons that the property of being naked is not limited to a stage of Tarzan’s life, but is intrinsic to the identity of this fictional character.

(28) ??\textit{Tarzan es desnudo}. 

19
‘Tarzan is \textit{is\textsubscript{SER}} naked.’

(Fábregas, 2012, p. 7)

For Fábregas, the IL/SL contrast gets at “the crucial question of which properties are considered essential and which properties are not, and how this reflects a deeper ontology of how humans categorize reality around them.” (2012, p. 9). As for relying on this distinction to explain the behaviour of \textit{ser} and \textit{estar}, despite the degree of overlap between the copulas and the IL/SL classification, it is apparent that there is not a direct mapping between them.

To conclude this discussion, there is much evidence to suggest that there is not a perfect one-to-one correspondence between the categorization of IL/SL predicates and the distribution of \textit{ser} and \textit{estar} with adjectival predicates. In the next section, I turn to studies that attribute the contrast between the copulas to a pragmatic distinction.

2.3.3 Pragmatics

Studies by Falk (1979), Franco and Steinmetz (1983, 1986), Clements (1988), Escandell-Vidal and Leonetti (2002), and Maienborn (2005) have focused on the dimension of pragmatics in copula selection. The question of the comparisons implied by copula constructions forms the basis for much of the pragmatic theory on \textit{ser} and \textit{estar}. Falk (1979) observes that both types of copula clauses express a comparison between the subject referent and a norm interiorized by the speaker. He suggests that the speaker, in choosing which copula to use with an adjective, must decide between the expression of a \textit{class norm} (\textit{ser}) and an \textit{individual norm} (\textit{estar}). An adjective that expresses a class norm classifies the subject as part of a group that shares the same characteristic. An adjective that expresses an individual norm compares the current state of the subject with a previous state and classifies it as a departure from the subject’s normal behaviour.

In a similar vein, Franco and Steinmetz (1983, 1986) propose a comparison analysis of \textit{ser} and \textit{estar} in which \textit{[ser + adjective]} expresses a comparison between X and Y, and \textit{[estar + adjective]} expresses a comparison between X and X. In the X/Y comparison, X is the subject of the
comparison and Y is the world and other individuals, which form the standard for the quality being compared; in the X/X comparison, the subject of the comparison is compared to itself at different times. The X/X comparison expressed by [estar + adjective] compares X at two stages in time; however, it can also express a comparison between X and the speaker’s anticipated idea of X, that is, what the speaker expected to encounter in X. Thus, there are two X/X comparative meanings that [estar + adjective] may convey: the first presupposes a prior experience that the speaker has had with the subject and is a comparison between the current state of the subject and the state observed by the speaker during prior experience; the second does not presuppose any prior experience with the subject by the speaker, but rather compares the speaker’s expectations with the current experience of the subject. Franco and Steinmetz demonstrate these meanings using the example sentences in (29).

(29)  
a.  \textit{Esta playa es buena.}  
\textquote{This beach is\textsubscript{SER} good/nice.}\quad \text{(1986, p. 381)}

b.  \textit{Esta playa está buena.}  
\textquote{This beach is\textsubscript{ESTAR} good/nice.}

In (29a), \textit{esta playa} is compared with other beaches in the world, and in (29b) \textit{esta playa} is compared with a previous conceptualization of the beach by the speaker, be it real or imagined. The authors also mention a third meaning for (29b), whereby the speaker selects \textit{estar} to express an observation about the subject that is limited to that moment in time, implying that the speaker reserves judgment on extending the observation beyond the present moment. In sum, [estar + adjective] may have up to three potential interpretations in this analysis: two kinds of X/X comparative meanings, and a spatio-temporally limited meaning, which also carries an implicature of comparison.

As in Franco and Steinmetz, the perspective of the speaker is also a factor in Clements’s (1988) pragmatic approach to the copula contrast. Building on Falk’s work on the class versus individual norm, but recasting the idea of norm in terms of frame of reference, he integrates
additional pragmatic factors into his analysis. Clements argues that the animacy of the subject and the perspective of the speaker are relevant to an analysis of copula selection with adjectival predicates. Another crucial element for Clements’s analysis is the presence or absence in the discourse context of the presupposition of a nexus or connection to another situation or state, encoded in the feature [+/-NEXUS]. *Estar* presupposes a nexus and corresponds to the feature [+NEXUS], whereas *ser* does not presuppose a nexus and thus corresponds to the feature [-NEXUS].

The notion of a [+/-NEXUS] feature is further developed by Maienborn (2005). She claims that the syntax and semantics of *ser* and *estar* are identical, and the only difference between them is that *estar* predications are restricted to a specific discourse situation by the presence of the [+NEXUS] feature. However, there are many distributional and interpretational differences between the copulas that the [+/-NEXUS] feature alone cannot account for. Maienborn’s analysis seems to make sense in the [copula + adjective] context, but runs into trouble when faced with syntactic contrasts between the copulas in the differences in the types of elements they combine with. To mention but a few examples, *ser* but not *estar* may appear with a nominal predicate, *estar* is ungrammatical (or at least odd) as a participle in the progressive (??Javier está estando amable, ‘Javier is being friendly’), their distribution with prepositional complements is largely complementary, and the two verbs give rise to distinct interpretations with participles in passives (see section 2.1 for the full discussion of the distribution of *ser* and *estar*). None of these syntactic facts can be accounted for by a presupposition at the level of pragmatics.

Escandell-Vidal and Leonetti (2002) also take a pragmatic approach to *ser* and *estar* in a theory that focuses on explaining contexts where IL predicates are interpreted as SL. They present evidence that with the right context, some IL adjectival predicates are acceptable with *estar* (30).

(30)  

a.  

*estar rojo* (cf. *ser rojo*)

‘to look red or reddish, to turn red’

b.  

*estar inteligente* (cf. *ser inteligente*)

‘to behave in an intelligent way’
These kinds of contexts are cases of coercion, according to the authors. Following previous studies on coercion (de Swart, 1998; Fernald, 1999; Kratzer, 1995), Escandell-Vidal and Leonetti define the concept as: “a reinterpretation process set up to eliminate the conflicts between the semantic content of a constituent and the requirements of other elements in the same construction.” (2002, p. 163). They give a coercion-based explanation for IL predications with estar that receive an SL reading (31).

(31) ¡Vaya! ¡Estás muy británico!

‘Wow! You look/are acting ESTAR British!’

In their application of the concept, coercion is a reinterpretation mechanism activated by a syntactic trigger, which in this case is estar, when followed by an IL predicate. In semantic composition a spatio-temporal variable is supplied to make up for the mismatch between a triggering element that requires such a variable (e.g., estar) and a constituent (e.g., an IL predicate) that lacks one.

A problem with Escandell-Vidal and Leonetti’s analysis is that the conditions that are considered to trigger coercion are defined rather broadly. A mechanism as powerful as coercion would need to be subject to further constraints in order to prevent its application in contexts where it would be unmotivated. Furthermore, Arche (2006) observes that Escandell-Vidal and Leonetti’s definition of coercion does not adequately capture the behaviour of many adjectives that regularly appear with estar. The estar cases in (32b) are not the outcome of a pragmatic reinterpretation process, and coercion is too strong a mechanism to apply to these contexts.

(32) a. Pablo es guapo/moreno/gracioso

‘Pablo isSER handsome/dark-skinned/funny’

b. Pablo está guapo/moreno/gracioso
‘Pablo looks/isESTAR handsome/tanned/funny’

(Arche, 2006, p. 239)

Arche suggests that it would be appropriate to refer to coercion in contexts like (33) or (34).

(33)  **Juan está muy americano (últimamente)**

‘Juan isESTAR behaving in the way Americans usually do.’

‘In his opinions, Juan is taking the Americans’ side.’

(34)  **Juan está muy democrático**

‘Juan isESTAR behaving very democratically.’

(Arche, 2006, pp. 251-252)

The predicates with *estar* in (32b) can be accounted for in terms of a change-of-state or aspectual analysis, but Escandell-Vidal and Leonetti’s coercion analysis should be reserved for cases such as the appearance of *estar* with denominal adjectives like those in (33-34).

I leave the discussion of pragmatic accounts of the copula distinction here. These proposals collectively emphasize the importance of the discourse context in the construction of meaning between the copula and an adjectival predicate. It is now time to revisit the question of aspect, and consider how recent developments in syntactic theory have been applied to solving the *ser/estar* puzzle.

### 2.3.4 An asymmetrical aspectual analysis

One of the central themes in the theoretical literature on *ser* and *estar* revolves around the question of the whether there is symmetry to be found in the structures of these verbs. In the accounts of Luján (1981) and Roby (2009), symmetry exists in the form of the [−PERFECTIVE]
feature contrast. Many authors express the intuition that some kind of parallel should be present in the syntax of the copulas, given the extent of their overlapping contexts. In this section, I consider a proposal that departs from this perspective, and offers an asymmetrical aspectual analysis for *ser* and *estar*.

Schmitt (1996, 2005) proposes an aspectual account of the contrast between *ser* and *estar*\(^3\), but with some key differences from previous approaches. To recap, aspect is conceived of elsewhere as a lexical property of verbs or as a functional head that belongs to the extended projection of a verb. However, rather than explaining aspect in terms of a [+/-PERFECTIVE] feature or an AspP, Schmitt takes a compositional approach to aspect and the IL versus SL predicate-type distinction. *Ser* is underspecified for aspect and lacks internal temporal or aspectual structure; *estar* is specified for aspectual properties and universally quantifies over states. *Estar* is syntactically more complex than *ser* and does not behave consistently as a copula in the classic sense. *Estar* is an aspectual verb that takes [+V] predicates, and *ser* is a true copula and takes [+N] predicates.

Schmitt (2005) draws upon the Distributed Morphology (DM) theory of morphosyntax (Halle & Marantz, 1994; Marantz, 1997) and the computational linguistics approach of Pustejovsky’s (1995) Generative Lexicon (GL). She adopts the structure in (35) for the verb phrase, including Kratzer’s Voice Phrase (1996) to introduce the subject of a copula clause.

\(3\) Schmitt’s focus is on the copular distinction in Brazilian Portuguese, which closely resembles, but is not isomorphic to, the Spanish dual copula system. The example sentences in this section are translated from Portuguese, and Schmitt’s analysis also holds for Spanish in the examples cited.
Following DM, roots combine with features in the syntax, including category features, to project a single head. A copula is either $v$ or $v+P$, where $P$ expresses an aspectual feature. Schmitt describes $ser$ as a “transparent verbalizer”, meaning that its structure is purely $v$: categorial information only with no $P$ features, hence its capacity to carry verbal morphology but no aspectual information.

In GL, verbs are composed of subevents, which can be states, processes or transitions. Copula verbs are composed of two subevents, and only combine with states. The first subevent is associated with the copula, and the second subevent with the complement. The head-complement relation of this structure ensures that the two subevents are predicated of the same argument in order to produce a single complex event structure (36).

(36) Copula and copula-like VP structure (Schmitt)

```
   vP
   / \        \xP
v(+P)     xP
```

Given that $ser$ has no aspectual information and does not denote any eventuality type, it has an atemporal reading by default when no other elements in the clause trigger an eventive or temporal reading. At the same time, this absence of event structure is what allows $ser$ the flexibility to appear in a wider variety of contexts than $estar$. $Ser$ is compatible with both stative and eventive predicates. $Estar$, which always denotes a state, has the structural representation $v + P[STATE]$. This limits the range of contexts where it can appear and explains why it does not allow for eventive readings\(^4\). $Estar$ predicates always have stative interpretations, according to Schmitt, because of $estar$’s $P$ features that denote a state type of subevent.

\(^4\) See section 2.4.2 for further discussion of the stative/eventive contrast.
An implied contrast is present in the dual copula system in Spanish and Portuguese. Schmitt suggests that the implication of temporariness that accompanies some *estar* predicates (38) is a simply a consequence of the fact that there is another option in the language (37).

(37)  
*Él es feliz.*  
(Translation from Portuguese)  
‘He is *SER* happy.’  
(Schmitt, 2005, p. 139)

(38)  
*Él está feliz.*  
(Translation from Portuguese)  
‘He is *ESTAR* happy (now).’  
(Schmitt, 2005, p. 139)

Schmitt’s analysis of the meanings and structures of copulas in Portuguese (and Spanish, by extension) represents an alternative to approaches that rely on a lexically-based analysis of aspect, or explain the contrast between *ser* and *estar* clauses in terms of IL versus SL predicates.

### 2.3.5 Prepositional explanations

Recently, a number of approaches have developed the idea that prepositions play a role in the composition of meaning in copular clauses (Arche, 2006, 2012; Brucart, 2012; Fábregas, 2012; Gallego & Uriagereka, 2009; Zagona, 2012). The debate over whether a structural parallel exists between *ser* and *estar* is one that has surfaced in various guises over the years – from classic studies, to aspectual accounts, to pragmatic explanations – and prepositional analyses are no exception. In the most recent incarnation of the polemic, Brucart (2012) and Fábregas (2012) pursue symmetrical approaches, whereas Zagona (2012) advances an asymmetrical proposal.

For Zagona, the *ser/estar* contrast is attributable to the presence or absence of a preposition. In her analysis, *estar* is a light verb with an uninterpretable preposition feature \([uP]\) that is valued by a matching feature in its complement. This complement can take the form of a PP or an AspP
– the latter, when *estar* acts as an auxiliary of the progressive. *Ser*, on the other hand, has no such feature, and thus is free to combine with a much wider range of complements.

The notion of a “path” component in copula semantics and its realization in the form of a preposition is developed in Brucart (2012). In the spirit of Gallego and Uriagereka’s (2009) proposal that *ser* merges with a preposition to produce *estar*, the core idea of Brucart’s analysis is that an interpretable feature of terminal coincidence accompanies *estar*, whereas *ser* is unmarked. His approach to copular clauses proposes a relation of coincidence between figure and ground to explain the “abstract path” meaning that these sentences convey. The figure is defined as the object, person, place or event that is located, and the ground is the location itself. Brucart adopts a symmetrical analysis of the copulas based on den Dikken’s (2006) Relator (R) phrase, but diverges from den Dikken’s original conception of an R head, which is realized as the copula in a relation with a small clause. Instead, he suggests that both *ser* and *estar* are merged in AspP above an R head in vP. The R head is spelled out as a coincidence preposition, either terminal or central. The basic idea of coincidence is “a spatial, temporal or identity relation between two elements, one functioning as a figure and the other being a ground” (Brucart, 2012, p. 17). In central coincidence, the figure coincides with the ground at the centre of the path; in terminal coincidence, the figure and the ground do not coincide at the centre of the path, and so the path either converges or diverges at the terminal point. *Estar* is related to a terminal coincidence preposition (this is responsible for the abstract path interpretation), whereas *ser* is unmarked but is usually related to a central coincidence preposition in the Relator phrase.

Fábregas (2012) endorses Brucart’s interpretation of the prepositional hypothesis and builds upon it further. He proposes that *estar* instantiates the terminal node of an aspectual projection containing a terminal coincidence relation, and that *ser* is used if there is no terminal preposition to check. The subject of PredP is a spatiotemporal variable in SL predicates, such as *estar desnudo* (39). In the case of IL predicates such as *ser inteligente*, it is an individual (40). The SL predicate structure encodes a predication relation between the adjective and a stage of an individual represented by the variable; the IL predicate structure encodes a predication relation between the adjective and the individual.
In the case of evidential uses of *estar*, Fábregas suggests that an AspP selects a PredP whose subject is an individual, and the aspectual [terminal] head relates the predication to a [pivot], which is the entity perceiving the characteristic property. The notion of a [pivot] comes from Sells (1987): “the entity with respect to whose perspective the content of a proposition is true” (Fábregas, 2012, p. 52). The [pivot] feature will undergo feature-checking with the CP domain when T and C merge, producing the interpretation in which a proposition is restricted by the perspective of [pivot]. The feature-checking by the pivot is illustrated in the structure in (41) for the sentence *la sopa está fenomenal* (‘the soup *iESTAR* wonderful’).
This concludes the discussion of proposals that have attempted to explain the contribution of the copulas *ser* and *estar* to the *[copula + adjective]* construction. In the next section, I consider the role of the adjective in determining the distribution and interpretation of *ser* and *estar* with adjectival predicates.

2.4 Properties of the adjective

In this section, I review some notable proposals regarding the nature and function of adjectives, and then evaluate several studies that relate these theories to Spanish. The two main approaches for the semantic classification of adjectives that I discuss include the eventive/stative contrast, and a gradability analysis.

2.4.1 Semantic classification of adjectives
Within the typological literature, Dixon’s (1977, 2004) work on adjective classes is foundational. Using cross-linguistic data and a series of morphosyntactic and semantic diagnostics, Dixon presents a system of classification of adjective types (42).

(42) a. Dimension (e.g., big, small, long, tall, wide, deep)

b. Age (e.g., new, young, old)

c. Value (e.g., good, bad, perfect, real, strange, important, necessary, lucky)

d. Colour

e. Physical property (e.g., hard, heavy, wet, strong, clean, hot, sour, and a subclass of corporeal properties like well, sick, tired, dead, absent)

f. Human propensity (e.g., jealous, happy, clever, cruel, eager)

g. Speed

h. Difficulty (e.g., easy, hard, simple)

i. Similarity (e.g., similar, different)

j. Qualification (e.g., definite, true, probable, common, appropriate)

k. Quantification (e.g., all/whole, many, few)

l. Position (e.g., high, low, near, far, right, left, northern, southern)

m. Cardinal numbers (also first, last)

(Dixon, 1977, p. 31)

Dixon observes that the behaviour of adjectives can be categorized as either “verb-like” or “non-verb-like” (2004, p. 14). On one hand, verb-like adjectives may appear in the absence of a verb, and with some or all of the verbal morphosyntax of a language. On the other hand, non-verb-like adjectives tend to appear as copula complements. If a language has both types of adjectives (e.g.,
Igbo), then verb-like adjectives express transient states and non-verb-like adjectives express more permanent properties (Dixon, 2004, p. 32). The typological data suggest that the expression of the contrasting meanings of property versus state is a widespread, if not universal, characteristic of adjectives across languages.

2.4.2 The stative/eventive contrast

Along the same lines as Dixon’s verb-like/non-verb-like distinction, one approach to adjective classification divides adjectives into stative versus eventive categories. Stative adjectives include stable properties, such as dimension, age, colour, and so on, whereas eventive adjectives encompass transitory conditions, such as human propensities. Eventive adjectives can enter into progressive (43a) and imperative (44a) contexts, as opposed to stative adjectives, which are ungrammatical in these contexts (43b, 44b).

(43)  
   a. Él está siendo travieso.  
       ‘He is\textit{Estar} being\textit{Ser} naughty.’
   b. #Ella está siendo alta.  
       ‘She is\textit{Estar} being\textit{Ser} tall.’

(44)  
   a. Sé amable.  
       ‘Be\textit{Ser} nice.’
   b. #Sé joven.  
       ‘Be\textit{Ser} young.’

\footnote{The examples feature the verb \textit{ser}, but \textit{estar} is accepted in very limited circumstances with imperatives (i) (with participle and perfective adjectives) and not at all in the progressive, to the best of my knowledge. 
(i) \textit{Estáte quieto/attento/preparado}.  
    ‘Be\textit{Estar} still/attentive/prepared.’}
Several studies have examined the consequences of the stative/eventive adjective contrast for copula selection in Spanish (Arche, 2006; Bosque, 1990; Fábregas, Leferman, & Marín, 2012). Bosque (1990) approaches the issue of aspectual interpretations of adjectives, and by extension the copulas with which they combine, through the lens of diachronic morphosyntax. He rejects the notion of a binary [+/- PERFECTIVE] feature for adjectives, contra Luján (1981). Instead, Bosque claims that adjectives that evolved from verbs with an eventive argument have a perfective reading (45), in contrast to other adjectives (46).

(45)  *lleno, suelto, limpio, seco*  
‘full’, ‘loose’, ‘clean’, ‘dry’

(46)  *bueno, alto, inteligente, elegante*  
‘good’, ‘tall’, ‘intelligent’, ‘elegant’

(1990, p. 178)

Perfective adjectives denote a state that results from an action or process, which suggests these adjectives have verbal roots. They are derived from participles through a morphological process he refers to as conversion or truncation, which preserves the eventive argument of their verbal root (47).

(47)  *llenar*  \(\rightarrow\)  *llenado*  \(\rightarrow\)  *lleno*  
‘to fill’ [INF]  \(\rightarrow\)  ‘filled’ [PART]  \(\rightarrow\)  ‘full’ [ADJ]

(1990, p. 186)

However, Bosque’s proposal does not address the fact that some perfective adjectives, such as *limpio* ‘clean’ and *seco* ‘dry’, also combine with *ser* but without the resultative interpretation (48).

(48)  *El desierto es seco.*  
‘The desert **is** SER dry.’
In combination with *ser* in (48), the adjective *seco* is synonymous with *árido* ‘arid’, which is a different sense of the word than the resultative meaning that arises when it combines with *estar* (49).

(49)  *Las toallas están secas.*

‘The towels are *ESTAR* dry.’

An analysis which relies on the presence of an eventive argument in the semantics of the adjective itself cannot account for the different senses available for perfective adjectives like *seco* and *limpio*.

Other approaches break down the components of the predicate structure differently, essentially separating the eventive element from the lexical contribution of the adjective. Arche (2006) claims that not all IL predicates are stative, but can be either stative or eventive. Adjectives denoting mental properties (aptitudes and human dispositions according to Dixon’s (1977) classifications) or velocity may produce eventive readings with *ser* (50).

(50)  a.  *Juan fue muy cruel con Pedro*

‘Juan was *SER* very cruel with Pedro.’

b.  *Juan era rápido/lento a propósito*

‘Juan was *SER* quick/slow on purpose.’

(Arche, 2006, p. 92)

Arche also looks to additional properties of the predicate (the presence of a PP complement) and the agentive properties of the subject to account for the eventive interpretation of *ser* in sentences like (50a). The subset of IL predicates that permit an eventive interpretation with *ser* only do so if the subject has agentive properties (50b).

Fábregas, Leferman, and Marín (2012) propose an analysis of the class of eventive IL adjectives which they refer to as evaluative adjectives (EAs). They describe EAs as a subset of adjectives that present mixed aspectual behaviour in both English and Spanish (51).
Fábregas et al. suggest that an event is involved in EAs. They propose that EAs are Davidsonian states, and treat them as eventive (but still IL) predicates. Crucially, the event is not part of the adjective itself, but rather is associated with the adjective through additional functional structure. This accounts for the exceptional behaviour of EAs without assigning any exceptional morphosyntax to the adjective itself. If the morphosyntax of EAs is no different from that of other IL adjectives, perhaps all IL adjectives are eligible to combine with an event through the proposed functional structure, and acceptability is determined pragmatically. However, the question of how pragmatic acceptability is determined in not addressed in this analysis.

2.4.3 Gradability

Bolinger (1972) is often cited as being among the first to describe the property of gradability, which subsequent studies have used to develop a set of criteria for the classification of adjectives. A basic diagnostic for distinguishing gradable (52) from non-gradable (53) adjectives is acceptability with degree adverbs such as very.

(52)  a.  *Ella es muy traviesa.*

   ‘She is very naughty.’

   b.  *Él es muy alto.*

   ‘He is very tall.’

(53)  a.  *#Ella es muy mexicana.*

   ‘She is very Mexican.’
b. #La bomba es muy atómica.

‘The bomb is very atomic.’

Non-gradable adjectives include adjectives denoting nationality (e.g. mexicano ‘Mexican’ in 53a) and other denominal adjectives (e.g. atómica ‘atomic’ in 53b). However, the gradable/non-gradable contrast is not isomorphic to the stative/eventive distinction, as the adjectives in (52) demonstrate.

Kennedy and McNally’s (2005) study represents an important advance in the development of a theory of the gradable properties of adjectives. The authors develop a proposal for the mechanisms of gradability, basing their analysis on the parameters of scale structure and standard of comparison. The scale structure of an adjective is lexically specified according to three parameters. The first is a set of degrees, which are points or intervals of measurement. The second is a dimension, which specifies the characteristic that is being measured. The third is an ordering relation, which can be positive or negative. A gradable adjective, then, denotes a relation between a degree of a dimension, and an individual. To consider a specific example, Kennedy and McNally provide the following denotation for the adjective expensive (54).

$$[[	ext{expensive}]] = \lambda d \forall x. \text{expensive}(x) = d$$

(2005, p. 349)

Scale structure can be open or closed. Open-scale predicates lack minimal and/or maximal elements (e.g., long, expensive, old), whereas closed-scale predicates have minimal and/or maximal elements (e.g., full, invisible, closed). The diagnostic for distinguishing between open- and closed-scale predicates is their ability to combine with proportional modifiers, such as half, mostly and most of the way (55).

a. #half long, #mostly expensive, #most of the way old

b. half full, mostly invisible, most of the way closed

(Kennedy & McNally, 2005, p. 352)
Closed-scale adjectives may be totally closed, where the scale has both a minimal and a maximal element, or partially closed, where the scale has either a minimal or a maximal element. In the complete typology, there are four categories (56).

(56)  
   a. totally open scale (e.g. tall/short, deep/shallow, eager/uneager)  
   b. lower closed scale (e.g. straight, quiet, unknown)  
   c. upper closed scale (e.g. certain, pure, safe)  
   d. totally closed scale (e.g. full/empty, open/closed, visible/invisible)

(Kennedy & McNally, 2005, p. 354)

The second parameter of gradability in adjectives is that of the standard of comparison. The standard of comparison may be defined as relative (contextually determined) or absolute (non-contextually determined), and it characterizes the truth conditions of the predicative use of scalar adjectives. In the case of an adjective with a relative standard of comparison, the scalar property may be true of an individual in one context and false in another. However, predicates whose standard of comparison is absolute are true of their arguments regardless of context. For instance, tall is an example of a scalar adjective with a relative standard of comparison. An individual who is considered tall within the context of his or her family may not appear tall in the context of a basketball team. An adjective like empty is a typical case of a scalar predicate with an absolute standard of comparison. All that matters, as far as the truth conditions for empty are concerned, is that the entity to which this predicate applies contains nothing. Scalar adjectives with totally open scale structure have relative standards of comparison; those with totally or partially closed scale structures are associated with absolute standards of comparison (57).

(57)  
   a. totally open scale: Michael is tall.  
   b. lower closed scale: The baby is awake.  
   c. upper closed scale: The door is closed.  
   d. totally closed scale: The glass is empty.
Scale structure and the standard of comparison are related in the semantic composition as follows. Kennedy and McNally adopt an analysis that makes use of a null degree morpheme, \textit{pos}, in the case of unmodified adjectival predicates. The function of \textit{pos} is to encode a relation between the degree argument of the adjective and the standard of comparison. They provide a denotation of the predicate \textit{expensive}, in which \textit{pos} composes with \textit{expensive} (58).

\begin{equation}
[[\text{pos}]][[[\text{expensive}]]) = \lambda x. \exists d (\text{standard}(d)([[\text{expensive}]])(C) \land [[\text{expensive}]](d)(x))
\end{equation}

Kennedy and McNally’s analysis, based on the parameters of scale structure and standard of comparison, serves as a cornerstone of Gumiel-Molina and Pérez-Jiménez’s (2012) approach to the \textit{ser/estar} puzzle.

In a somewhat different approach to scale structure, Husband (2012) indirectly addresses many of the recurring questions that surface in studies of \textit{ser} and \textit{estar} in his work on the aspectual interpretation of states. Contra Kennedy and McNally (2005), who propose that the components of gradability are part of the lexical semantics of adjectives, Husband locates the elements and composition of gradability in the functional structure. He draws a parallel between telicity in events and the existential interpretation of subjects in states, and suggests a unified theory of quantization in events and states. Arguing that the same mechanisms of quantization apply across multiple domains, the structures that enable individuation and counting for events, nominals, scales, and states come about via the same compositional mechanisms (see Table 2.2).
Table 2.2. Quantized structures (Husband, 2012, p. 151)

<table>
<thead>
<tr>
<th></th>
<th>Homogeneous</th>
<th>Quantized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events</td>
<td>atelic</td>
<td>telic</td>
</tr>
<tr>
<td>Nominals</td>
<td>unspecified quantity</td>
<td>specified quantity</td>
</tr>
<tr>
<td>Scales</td>
<td>open scale</td>
<td>closed scale</td>
</tr>
<tr>
<td>States</td>
<td>individual-level</td>
<td>stage-level</td>
</tr>
</tbody>
</table>

Following Borer’s (2005) definition, a homogeneous predicate is one that is both cumulative and divisive (59).

(59) **Homogeneous**: $P$ is homogeneous iff $P$ is cumulative and divisive.

a. $P$ is *cumulative* iff $\forall x, y[P(x) \& P(y) \Rightarrow P(x \cup y)]$

   $P$ is *cumulative* iff for all $x$ and $y$ with property $P$, the union of $x$ and $y$ also has property $P$.

b. $P$ is *divisive* iff $\forall x[P(x) \Rightarrow \exists y[P(y) \& y < x] \& \forall x, y[P(x) \& P(y) \Rightarrow P(y) \& y < x \Rightarrow P(x - y))]$

   $P$ is *divisive* iff for all $x$ with property $P$ there is a proper part $y$ of $x$ which also has property $P$, and for all $x$ and $y$ with property $P$ if $y$ is a proper part of $x$ then the subtraction of $y$ from $x$ also has property $P$.

   (Husband, 2012, p. 74)

If either of these properties is not present, the predicate is interpreted as quantized. The homogeneous versus quantized interpretation is determined compositionally in the syntax. The quantization of the state is the result of the mapping from the scale structure of the adjective to the part structure of the state. Husband proposes that stative predicates introduce their external arguments by means of a VoiceP (Kratzer 1996), and presents evidence that VoiceP is present
for stative verbs as well as eventive verbs. In this way, VoiceP maps the part structure of the state to its external argument. This approach contrasts with theories based on information structure, which correlate strong subjects with existential interpretations owing to topic requirements.

Husband’s theory of quantization can be called upon to explain the interpretations of *ser* and *estar* with adjectival predicates. Here, quantization is what determines the availability of an existential interpretation of a subject. Specifically, homogeneous subjects do not license an existential interpretation; quantized subjects do. It is the stative Voice head that introduces the external argument and maps the part structure of the state to the part structure of the individual. The mediation of the part structure of the subject by VoiceP distinguishes Husband’s approach from other theories of the copula contrast, such as the IL/SL predicate-type distinction.

In an application of the predictions of studies of the gradability of adjectives (e.g., Husband, 2012; Kennedy & McNally, 2005), Gumiel-Molina and Pérez-Jiménez (2012) give an account of the distribution of *ser* and *estar* with adjectival predicates. Their central claim is that open-scale/relative adjectives appear with *ser* (60b) and closed-scale/absolute adjectives appear with *estar* (60a).

(60)  

(a)  *La pared está {seca/húmeda/recta}*  

‘The wall is*ESTAR {dry/wet/straight}.’  

(b)  *Juan es {inteligente/alto/delgado}*  

‘Juan is*SER {intelligent/tall/thin}.’

(Gumiel-Molina & Pérez-Jiménez, 2012, p. 45)

Gumiel-Molina and Pérez-Jiménez follow Husband in arguing that scalar properties are introduced in the syntax by functional structure. In the same way as the part-structure of a verb’s internal argument determines its telicity, the scale structure of an adjective determines the aspectual properties of the copular predicate. Hence, the copulas express a distinction between homogeneous and quantized states, parallel to the atelic/telic contrast in the eventive domain.
Different types of pos morphemes generate differing entailment patterns with absolute/closed-scale and relative/open-scale adjectives (61).

\[(61)\]

\(\text{a. } [[\text{pos open}]] = \lambda g \in D_{<e,d>} \lambda x \ [(g(x)) \geq \text{std}(g)]\)

\(\text{b. } [[\text{pos lower-closed}]] = \lambda g \in D_{<e,d>} \lambda x \ [(g(x)) > \text{min}(g)]\)

\(\text{c. } [[\text{pos upper-closed}]] = \lambda g \in D_{<e,d>} \lambda x \ [(g(x)) = \text{max}(g)]\)

(Gumiel-Molina & Pérez-Jiménez, 2012, p. 53)

The truth conditions for a maximum-standard absolute adjective entail that its argument has a maximal amount of the property expressed by the adjective (61c). This entails that no other individual may exceed it in this property. (The inverse also applies for minimum-standard adjectives (61b).) On the other hand, for a relative adjective, what is relevant is the context, so its argument is not held to a minimum or maximum standard of comparison, but a contextually-determined one (61a). In comparative constructions, absolute adjectives give rise to positive and negative entailments, as per the adjective’s polarity. Relative adjectives do not generate contradictions when followed by an assertion of its polar opposite property. Gumiel-Molina and Pérez-Jiménez demonstrate these diagnostics in Spanish and the corresponding distributional patterns of ser and estar with the minimum-standard adjective húmedo ‘wet’ (62a), the maximum-standard adjective seco ‘dry’ (62b), and the relative adjectives alto ‘tall’, delgado ‘thin’ and inteligente ‘intelligent’ (63).

\[(62)\]

\(\text{a. }\) Minimum-standard adjectives: X is more ADJ than Y \(\Rightarrow\) X is ADJ

\(\text{El suelo está más húmedo que la encimera } \Rightarrow\text{ El suelo está húmedo}\)

\(\text{‘The floor is}_{\text{ESTAR}}\text{ wetter than the counter.’ } \Rightarrow\text{ ‘The floor is}_{\text{ESTAR}}\text{ wet.’}\)

\(\text{b. }\) Maximum-standard adjectives: X is more ADJ than Y \(\Rightarrow\) Y is not ADJ

\(\text{El suelo está más seco que la encimera } \Rightarrow\text{ La encimera no está seca}\)
‘The floor is\textsc{estar} drier than the counter.’ \textarrow \textbullet ‘The counter is\textsc{estar} not dry.’

(63) Relative adjectives: $X$ is more ADJ than $Y$ \textarrow $X/Y$ is (not) ADJ

\begin{quote}
\textit{Juan es más \{alto/delgado/inteligente\} que Pedro \textarrow \{Juan/Pedro\} (no) es}

\{alto/delgado/inteligente\}
\end{quote}

‘Juan is\textsc{ser} \{taller/thinner/more intelligent\} than Pedro.’ \textarrow ‘\{Juan/Pedro\} is\textsc{ser} (not) \{tall/thin/intelligent\}.’

(Gumiel-Molina & Pérez-Jiménez, 2012, p. 46)

The variable behaviour of many adjectives in Spanish is explained in terms of their combination with different types of pos morphemes. Adjectives which can combine with both \textsc{ser} and \textsc{estar} show contrasting entailment patterns with each copula. For instance, the adjective \textit{seco} ‘dry’ behaves as an absolute adjective with \textsc{estar} in (64), and as a relative adjective with \textsc{ser} in (65), according to entailment pattern diagnostics.

(64) \textit{El suelo está más seco que la encimera \textarrow La encimera no está seca}

‘The floor is\textsc{estar} drier than the counter.’ \textarrow ‘The counter is\textsc{estar} not dry.’

(65) \textit{El clima de Lisboa es más seco que el de Barcelona, pero ninguno de los dos es seco.}

‘The climate of Lisbon is\textsc{ser} drier than the climate of Barcelona, but neither of them is\textsc{ser} dry.’

(Gumiel-Molina & Pérez-Jiménez, 2012, p. 50)

The proposal mentions several other examples of adjectives that produce contrasting relative/absolute entailment patterns in combination with \textsc{ser/estar}, including \textit{bueno} ‘good’, \textit{malo}}

Gumiel-Molina and Pérez-Jiménez argue against accounts of the copula as an auxiliary or light verb, and suggest that *ser* and *estar* are both lexical verbs with the same structural representation. They provide the following structure for the syntax of predicative copular constructions (66) and the semantic denotation in (67).

(66) Predicative copular clause structure (Gumiel-Molina & Pérez-Jiménez)

```
VoiceP
  Voice
    DP
    Voice’
      Voice
        VP
          Vcopula
            ser/estar
              pos
                A
```

(67) \[[\text{be} _{\text{pred}}]\] = \lambda P \lambda x [P(x)]

(2012, p. 54)

An important prediction of this approach is that any adjective is free to combine with any pos morpheme, and, extending this to Spanish, any adjective should be able to combine with either copula. This prediction is not supported by the facts of the distribution of the copulas with adjectival predicates (see section 2.1). Nonetheless, in adopting elements of Kennedy and McNally’s analysis of gradability and Husband’s analysis of the quantization of states, Gumiel-Molina and Pérez-Jiménez’s proposal provides a convincing compositional account of gradability in [copula + adjective] contexts in Spanish.

2.4.4 Summary: Properties of the adjective
In this section, I have reviewed some of the theoretical literature on adjectives in general and adjectival predicates of Spanish copula clauses in particular. Certain types of adjectives engage in eventive behaviour with *ser* under specific circumstances, and these interpretations are attributed to functional structure rather than the lexical semantics of the adjectives themselves. I have also evaluated several studies of gradability, a topic which has been a major focus of recent theory on the copula contrast in Spanish. These proposals offer a promising account for the distribution of *ser* and *estar* with adjectives in terms of the quantization of states. In the next section, I turn to the background literature dealing with the context that is at the heart of this inquiry: the evidential use of *estar*.

### 2.5 The evidential use of *estar*

The evidential use of *estar* is the [estar + adjective] context that engages the viewpoint of the speaker, as illustrated by the sentences in (2), repeated here.

\[(2) \quad \begin{align*}
\text{a.} & \quad \text{Este vino está riquísimo.} \\
& \quad \text{‘This wine is} \text{ESTAR delicious}_{\text{INT}}.\text{’} \\
\text{b.} & \quad \text{¿Viste la nueva película de Almodóvar? Está buena.} \\
& \quad \text{‘Did you see the new Almodovar film? It is} \text{ESTAR good.’} \\
\text{c.} & \quad \text{Están muy lindos los cachorros del vecino.} \\
& \quad \text{‘The neighbor’s puppies are} \text{ESTAR cute.’}
\end{align*}\]

Although these uses of *estar* have mostly been overlooked in the literature on *ser* and *estar* with adjectives, several recent proposals address them in their analyses (Camacho, 2012a, 2012b; Escandell-Vidal & Leonetti, 2002; Roby, 2009).
2.5.1 Evidential coercion with *estar*

Perhaps the origin of the label *evidential* for the types of [*estar* + adjective] clauses that encode the perspective of the speaker can be traced to Escandell-Vidal and Leonetti (2002). These authors connect the concept of evidentiality to certain *estar* predications, in applying Fernald’s (1999) notion of “evidential coercion” to explain cases where an IL predicate appears with *estar*. These are sentences that give rise to a “behavioural” reading of the property denoted by the adjective, which is given an evidential interpretation by the speaker (Escandell-Vidal & Leonetti, 2002, p. 165). The sentences in (68) illustrate that these *estar* predications express the speaker’s perceptions and are “widely used to express judgments on taste and personal evaluation” (2002, p. 167).

(68)  
a.  *Esto está buenísimo.*  
    ‘This is *ESTAR* tastes really good.’  
b.  *Tu trabajo está bastante flojo.*  
    ‘Your work is *ESTAR* really poor.’  
c.  *John Goodman está genial en esa película.*  
    ‘John Goodman is *ESTAR* great in that movie.’

(Escandell-Vidal & Leonetti, 2002, p. 167)

For Escandell-Vidal and Leonetti, the evidential use of *estar* is subject to few constraints. Clauses whose predicates involve epistemic adjectives like *necesario* ‘necessary’, *falso* ‘false’, and *evidente* ‘evident’ are excluded; the subjects of these clauses are propositional arguments rather than entities that can be perceived by the speaker. Otherwise, they say, “the limitations on the coercion process induced by *estar* are to be related to the conceptual cost of the operation involved, but not to any lexical idiosyncrasy or to any syntactic constraint” (2002, p. 168).
The main criticism of this approach is that coercion is a type-shifting operation which serves to reconcile a discrepancy between semantic types in the composition, but is too strong a mechanism to apply to the majority of cases of the evidential use of \textit{estar}, such as those in (68). To be clear, there are a few contexts where Escandell-Vidal and Leonetti’s coercion analysis genuinely applies, such as the contexts in (69-70), where the predicate is a denominal adjective.

(69) \textit{Juan está muy americano (últimamente)}

‘Juan is\textit{estar} behaving in the way Americans usually do.’

‘In his opinions, Juan is taking the Americans’ side.’

(70) \textit{Juan está muy democrático}

‘Juan is\textit{estar} behaving very democratically.’

(Arche, 2006, p. 251-252)

An explanation of the meaning differences between different types of \textit{[estar + adjective]} clauses, and in particular those with evidential meaning, must take into account properties of the adjective.

2.5.2 The discovery interpretation of \textit{estar}

Roby (2009) also mentions the evidential use of \textit{estar} in his proposal. He states that the evidential use of \textit{estar} is employed to “describe something that the speaker is perceptually experiencing for the first time” (2009, p. 17). The prototypical example of this appears in descriptions of the sensation of taste, where he notes that \textit{estar} is the preferred copula (71).

(71) \textit{Este jamón serrano está fenomenal.}

‘This Serrano ham is\textit{estar}/tastes phenomenal.’
This observation dates back to Querido’s (1976) discussion of copula choice in the context of a first sensorial experience. Querido imagines a scenario in which a botanist in the Amazon comes across an unfamiliar species of tree, which has yellow leaves. In reporting his finding, he is faced with a choice between *ser* and *estar* to describe the colour of the leaves (72).

(72)  

a. *Las hojas de este árbol son amarillas.* (Translation from Portuguese)

‘The leaves of this tree are*SER* yellow.’

b. *Las hojas de este árbol están amarillas.* (Translation from Portuguese)

‘The leaves of this tree are*ESTAR* yellow.’

(Querido, 1976, p. 354)

Given that the botanist has not had the opportunity to observe the tree throughout a full cycle of the seasons, he does not know if the colour of the leaves is consistent year-round or if it changes at different times of year. In other words, the standard criteria for copula selection do not apply in this situation. Querido suggests that the botanist report his observation using *estar*, because “*estar* is the appropriate copula to report a first sensorial experience by means of a predicate AP, without making any induction or generalization” (1976, p. 354). Querido’s description of the use of *estar* to express an initial sensory impression also features centrally in Maienborn’s (2005) analysis, and she dubs it the “discovery interpretation of *estar*” (2005, p. 160).

Roby expands upon and refines this definition by making two important additional observations. First, the evidential use of *estar* is not limited to first sensorial experiences, but may be employed in contexts of recalled experience as well (73).

(73)  

a. *Estos tacos están buenos.*

‘These tacos are*ESTAR* taste good.’

b. *Esos tacos en Laredo están buenos.*
‘Those tacos in Laredo are ESTAR/taste good.’

(Roby, 2009, p. 64)

Second, the evidential use of *estar* is optional, not obligatory, in contexts of immediate (or recalled) experience. Just as Querido’s hypothetical botanist was confronted with a dilemma in deciding which copula would describe the colour of the leaves most accurately, speakers are faced with a choice between *ser* and *estar* in evidential contexts. They are by no means required to select *estar*, Roby notes:

*Ser* may also be used instead of *estar* to describe something based on immediate evidence, especially if its usual state can be assumed. The use of *estar* for describing a given sensorial experience, initial, subsequent, or recalled, also tends to carry more emphatic weight than *ser*. (Roby, 2009, p. 65)

The observations added by Roby about the optionality of the evidential use of *estar* and its applicability to contexts of indirect as well as direct evidence have contributed to a clearer characterization of the phenomenon.

2.5.3 Implied comparisons with *estar*

The implicatures of comparison that can accompany *ser*/*estar* predications form the basis for other explanations of the evidential use of *estar*. For Camacho (2012a), the core notion in the semantics of *estar* predications is that of an event boundary. It is the inception of a state that *estar* expresses, and Camacho formalizes this as an inchoative feature. He suggests that the evidential use of *estar* can be accounted for if *estar* is associated with a beginning boundary rather than a stage. This allows for the possibility of a pragmatic comparison between the subject of an *estar* predication (74) and similar entities in other situations (Camacho, 2012a, p. 463).

(74)  *Este jamón está fenomenal.*

‘This ham is ESTAR phenomenal.’
Thus, the description of the ham in (74) does not express a stage, but rather an implicit comparison with the speaker’s expectation of the ham, based on general experience. Camacho’s approach integrates a syntactic explanation for the general distribution of [copula + adjective] contexts with a pragmatic proposal of implied comparison to account for the evidential use of *estar*.

The implied-comparison analysis Camacho adopts is similar to the one put forward by Franco and Steinmetz (1983, 1986) in which *estar* potentially enters into two types of X/X comparisons and a third, non-comparative reading (see section 2.3.3). The first type of comparison involves a comparison by the speaker between the current state of the subject of an *estar* predication and a previous experience of the subject by the speaker. The second type of X/X comparison entails a comparison by the speaker between the current state of the subject of an *estar* predication and a previously held mental representation of that subject by the speaker. It is this type of implicit comparison that Camacho refers to in his explanation of the evidential use of *estar*. The third type of interpretation that Franco and Steinmetz propose for *estar* predications is one in which the speaker does not engage in comparisons of the subject with itself or with the speaker’s previous expectations, but rather expresses a temporally-limited observation through the choice of *estar* over *ser*. This is similar to Querido’s (1976) reasoning in regards to the botanist’s dilemma as to how to best to represent his discovery of a new species of tree in the Amazon. Querido argues that *estar* best expresses the meaning that the leaves of the tree are observed to be yellow at the present moment, but without implying an extension of this observation any further in time. Unlike Franco and Steinmetz, and Camacho, Querido does not distinguish between this interpretation and the use of *estar* to express a first sensorial experience. In Querido’s approach, the speaker is not comparing the colour of the leaves to a previously held mental image of the leaves, but rather is responding to direct evidence only. The distinction between these two approaches (the “discovery interpretation”, to use Maienborn’s phrase, versus the implied comparison interpretation) raises the question of whether or not the evidential use of *estar* involves a prior expectation in the mind of the speaker. If the speaker indeed selects *estar* in the presence of an implicit comparison with an anticipated state of the subject, then it may be more accurate to describe these uses of *estar* as expressing mirativity rather than evidentiality.
In other work, Camacho (2012b) continues to build on the concept of an implied comparison in accounting for the evidential use of *estar*. He draws a parallel between the Tibetan evidential/aspectual morpheme ‘*dug*’ and *estar* predications with evidential readings. In this analysis, he adopts the notion of a comparison class as proposed in Gumiel-Molina and Pérez-Jiménez’s (2012) work on gradability in predicates of *ser* and *estar* (see section 2.4.3 for discussion of Gumiel-Molina & Pérez-Jiménez’s proposal). The gradability hypothesis associates adjectives with absolute interpretations with *estar*. An absolute interpretation entails a comparison between an individual and itself. By contrast, adjectives with relative standards of comparison involve comparisons between an individual and other individuals in a given context, and these predicates appear with *ser*. The “within-individual comparison” meaning in the *estar* predication is contextually located through the scope relations of the assertion structure. This allows for the possibility of an evidential reading.

Camacho reports that ‘*dug*’ is similar to *estar* in having aspectual meaning that is described as stage-level, but its primary semantic content is evidential in nature. So, whereas evidential readings are optional with *estar*, they are obligatory with ‘*dug*’. The obligatory evidential meaning of ‘*dug*’ arises from its interaction with the grammatical category of viewpoint in Tibetan. There are two types of viewpoint, *self* and *other*; viewpoint is an independent property of Tibetan that is absent in Spanish. In denoting location, viewpoint provides the necessary contextual location to license the evidential morpheme ‘*dug*’. Because Spanish lacks the grammatical viewpoint category, location must be established pragmatically, and evidential readings are thus optional. In support of this analysis, Camacho notes that deictic references amplify evidential interpretations with *estar* by lending further contextual anchoring to the location established by the clause structure of a thetic judgment. The evidential reading of (75a) is more salient than that of (75b) due to the contextual anchoring effect of the demonstrative.

(75)  

a. *Este jamón serrano está fenomenal.*

‘This Serrano ham *is* wonderful.’

b. *El jamón serrano está fenomenal.*

‘Serrano ham *is* wonderful.’
In addition to the observation that deictic markers enhance evidential readings of *estar* predications, Camacho notes that evaluative adjectives appear most often in the evidential use of *estar*. In the sentence in (76a), the adjective *abierta* ‘open’ does not produce an evidential reading with *estar*; in the sentence in (76b), the evaluative adjective *flojo* ‘weak’ gives rise to an evidential interpretation with *estar*.

(76) a. Cuando Juan llegó a su casa, la puerta estaba abierta.
   ‘When Juan arrived at home, the door was *estar* open.’

   b. Me dijeron que el trabajo estaba flojo.
   ‘They told me that the paper was *estar* weak.’

Camacho’s observations about the reinforcing effect of deictic references and the tendency for evaluative adjectives to appear in evidential readings of *estar* predications provide further insight into the evidential use of *estar*.

2.5.4 Summary: The evidential use of *estar*

The challenge for explanations of the evidential use of *estar* lies in formalizing the evidential component, or the viewpoint of the speaker, in a way that is compatible with a convincing overall proposal for the syntax and semantics of the copulas. For some authors, the viewpoint of the speaker is mediated by the pragmatics. For others, the semantic composition is what determines evidential readings. For others still, evidential meaning is encoded in the syntax. Accounting for the evidential use of *estar* represents an opportunity to explore questions about the mapping between elements of syntax-semantics and discourse-pragmatics. In this regard, the study of the evidential use of *estar* has broader implications for the field of linguistic theory as a
whole. In the next section, I turn to the background literature relating to evidentiality, and consider how it applies to the evidential use of *estar*.

## 2.6 Evidentiality

### 2.6.1 Definitions of evidentiality

In this section, I outline some of the questions and contributions from both the theoretical and typological literature on evidentiality pertinent to the study of the evidential use of *estar*. Evidentiality is generally taken to refer to the encoding of the source of information for a proposition. Since Chafe and Nichols’ (1986) groundbreaking anthology of typological and theoretical work on this topic, there has been much progress in the study of evidentials. Further data on evidentials in a wide range of languages has been collected, and a great variety of theoretical proposals have attempted to explain the facts. Many outstanding questions remain, however, and highlight the challenge of finding more granular approaches to this topic.

Anderson’s (1986) definition of the grammatical phenomenon of evidentiality includes the following conditions (77).

(77) a. Evidentials show the kind of justification for a factual claim which is available to the person making that claim, whether

- direct evidence plus observation (no inference needed)
- evidence plus inference
- inference (evidence unspecified)
- reasoned expectation from logic and other facts

and whether the evidence is auditory, or visual, etc.
Evidentials are not themselves the main predication of the clause, but are rather a specification added to a factual claim about something else.

Evidentials have the indication of evidence as in (a) as their primary meaning, not only as a pragmatic inference.

Morphologically, evidentials are inflections, clitics, or other free syntactic elements (not compounds or derivational forms).

(Anderson, 1986, pp. 274-275)

Each condition in this definition touches upon a key debate in the literature on evidentials. The condition in (77a) contains a rough description of the categories or types of evidence that are possible. The issue of the inventory of evidential types has been approached from a language-specific point of view in some studies, and with the goal of establishing a universal typology of evidentiality in others. The condition in (77b) is concerned with whether evidentiality is associated with the illocutionary or the propositional level of meaning. That is, do evidentials have something to say about the truth conditions of the main proposition, or do they contribute to the speech act instead? Recent proposals suggest that they may do either one. The condition in (77c) stipulates that an explanation of evidential meaning is to be found in the syntax/semantics, and not in the pragmatics of a clause that expresses evidentiality. The last condition (77d) is a definition of the kinds of morphological elements that may be evidentials. There is disagreement on the question of whether a definition of evidentiality should necessarily include morphological criteria. Below, I argue against the inclusion of morphological criteria in a definition of evidentiality for several reasons.

Due to its restrictive nature, Anderson’s definition is an example of what has come to be known as the “narrow” conceptualization of evidentiality. By contrast, in a neighbouring article from the same volume, Chafe (1986) uses the term evidentiality in what is known as the “broad” sense. In the broad sense, evidential meaning not only includes the type of evidence a speaker has for a proposition, but may also extend to the expression of the degree of reliability, probability or certainty of the proposition, according to the speaker. The class of expressions that qualify as
evidentials encompasses both lexical and grammatical elements. These divergent definitions of evidentiality have generated a range of points of view, which I summarize in this section.

2.6.2 Evidential morphosyntax

Cross-linguistically, evidentials are heterogeneous in both form and structure. Languages that encode evidentiality through grammatical means (e.g., Quechua, see Faller, 2002, among others) may express evidential meaning through dedicated inflectional paradigm or specialized markers, such as verbal affixes, clitics, particles, auxiliaries or copulas. Aikhenvald (2004) states that about a quarter of the world’s languages have grammaticalized evidentiality. Other languages (e.g., English) refer to source of information through lexical means, in the form of adverbs, modals or parenthetical verbs (Rooryck, 2001a, 2001b). In still other languages, semantic extensions may develop on existing grammatical forms in order to encode evidentiality. Aikhenvald (2004) uses the term “evidentiality strategies” to describe cases whereby an existing grammatical category acquires an evidential extension in meaning. Many categories and forms can develop the semantic features of evidentials, including non-indicative moods and modalities, future, past tenses, resultatives and perfects, passives, nominalization, complementation, person-marking and demonstratives (Aikhenvald 2004). Over time, evidentiality strategies may develop into grammaticalized evidentials. For example, the perfect tenses in Andean varieties of Spanish in contact with Quechua and Aymara have developed evidential meaning extensions (de Granda, 1994; Escobar, 1997; Klee & Ocampo, 1995; Sánchez, 2004). The extension of contrasts in the tense system to encode evidential meaning has also been observed in Paraguayan Spanish in contact with Guaraní (Palacios, 2008).

Should a definition of evidentiality include morphological criteria? The range of variation in how evidentiality is expressed cross-linguistically is much wider than a narrow definition would acknowledge. The fact that evidentiality strategies are a source for evidential systems demonstrates that the boundary between the lexical and the grammatical is not fixed. Furthermore, the notion of the obligatory nature of evidential marking in languages in which evidentiality is expressed by a morphological paradigm has recently been called into question.
Quechua is a language that has traditionally been described in these terms, but Faller (2002) makes clear that the language allows for sentences without evidentials. Many other such languages do not in fact show obligatory evidential marking, and current literature in the field suggests that the expression of evidentiality is almost universally optional. Taken together, these observations are problematic for the narrow definition of evidentiality, and suggest that the inclusion of morphological criteria in a definition of evidentiality is unnecessary.

Accounting for the diverse morphosyntax that evidentials display across languages has driven much recent scholarship in the field of evidentials. A proposal that is promising in its flexibility and capacity to capture many of the facts is Blain and Dёchaine’s (2007) Evidential Domain Hypothesis. Their work is firmly situated in the tradition of the broad definition of evidentiality, in the sense that evidentials are characterized as markers of the viewpoint of the speaker. The authors argue that evidentials can be introduced into the clause in a number of different positions. Depending on the type of viewpoint that the speaker expresses, evidentials differ as to the level of the clause where they are integrated. The four domains include CP (the domain of clause type), TP (the temporal domain), AspP (the aspektual domain) and vP (the predicate domain). The position of the evidential in the clause determines the semantic mapping of the speaker’s perspective. CP-external evidentials have illocutionary force, IP-external evidentials have temporal/modal force, and so on. They present data from Plains Cree and Cree Montagnais/Naskapi, which show a range of variation in how evidential meaning is mapped onto CP and IP domains. The proposal is applicable to data from other languages with different types of evidential marking.

This approach contrasts with proposals that evidentials are syntactic heads that occupy a fixed position in the functional hierarchy (e.g., Cinque, 1999; Rooryck, 2001a, 2001b; Speas, 2004). Cinque (1999) argues for a universal hierarchy of functional projections. He proposes that clause structure is made up of a sequence of functional heads, of which the left periphery comprises four speaker-oriented mood/modality categories (78).

(78) \[ \text{Mood}_{\text{Speech Act}} \rightarrow \text{Mood}_{\text{Evaluative}} \rightarrow \text{Mood}_{\text{Evidential}} \rightarrow \text{Mood}_{\text{Epistemic}} \]

(Cinque, 1999, pp. 90, 106)
The differences among these categories are extremely fine-grained, and some lexical items may be able instantiate multiple categories, especially those that are structurally adjacent. However, the co-occurrence of multiple evidentials in the same clause (as is the case in Plains Cree and Cree Montagnais/Naskapi, according to Blaine and Déchaine) is troublesome for theories that assume a unique syntactic position for evidentials. The distributed approach of the Evidential Domain Hypothesis captures more of the facts, given the heterogeneity of evidential morphosyntax across languages.

2.6.3 The internal structure of evidential systems

The definition of the semantic domain of evidentiality and its various subdomains is a central topic in the literature. In general, proposals for the internal structure of evidential systems tend to fall into one of three types: categorical typologies, implicational hierarchies or scales (Faller, 2002). Categorical typologies take a taxonomical approach to classifying the types of evidence and the evidential systems of languages. Implicational hierarchies predict the presence or absence of an evidential type based on the presence or absence of another element in the system. Evidential scales locate evidence types on a scale according to degrees of strength.

Typologies are by far the most common model for the structure of evidential systems. Cross-linguistic studies have identified the tripartite division of direct/inferred/reported evidence as the main types of source of information that are encoded by evidentials (see Aikhenvald, 2004; Faller, 2002; Willett, 1988; among others). The notion of evidence-type is the basis for the concept of an evidential in typological proposals, such as Willett’s (1988) taxonomy, presented in Table 2.3.
Table 2.3. Types of evidence (Willett, 1988, p. 57)

<table>
<thead>
<tr>
<th>Types of evidence</th>
<th>Direct</th>
<th>Indirect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attested</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual</td>
<td>Auditory</td>
<td>Other sensory</td>
</tr>
<tr>
<td>Second-hand</td>
<td>Third-hand</td>
<td></td>
</tr>
<tr>
<td>(hearsay)</td>
<td>Folklore</td>
<td></td>
</tr>
<tr>
<td>Reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferring</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reasoning</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Willett’s system contains three levels of classification. The first level distinguishes between the expression of direct or indirect evidence, and the second level divides evidentials into attested, reported and inferring types. At the third level are subcategories of attested, reported and inferring evidentials. Speas (2004, 2008) proposes a more limited set of universal evidential meanings based on the notion of evidence type. Her typology comprises the four categories of personal experience (internal sensation, experience), direct evidence (witnessing), indirect evidence (inference, assumption) and hearsay (quote). Aikhenvald’s (2004) system allows for languages to distinguish between two and five evidential types from an inventory of six possible parameters (Table 2.4).
Table 2.4. Semantic parameters in evidentiality systems (Aikhenvald, 2004, p. 65)

<table>
<thead>
<tr>
<th></th>
<th>I. Visual</th>
<th>II. Sensory</th>
<th>III. Inference</th>
<th>IV. Assumption</th>
<th>V. Hearsay</th>
<th>VI. Quotative</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 choices</td>
<td>firsthand</td>
<td></td>
<td>non-firsthand</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>firsthand</td>
<td></td>
<td>non-firsthand</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>firsthand</td>
<td></td>
<td>non-firsthand</td>
<td>different system or &lt;no term&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;no term&gt;</td>
<td></td>
<td>&lt;no term&gt;</td>
<td>reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 choices</td>
<td>direct</td>
<td></td>
<td>inferred</td>
<td>reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>visual</td>
<td></td>
<td>inferred</td>
<td>&lt;no term&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>visual</td>
<td></td>
<td>inferred</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>visual</td>
<td></td>
<td>inferred</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;no term&gt;</td>
<td></td>
<td>inferred</td>
<td>reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 choices</td>
<td>visual</td>
<td></td>
<td>inferred</td>
<td>reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>direct</td>
<td></td>
<td>inferred</td>
<td>assumed</td>
<td>reported</td>
<td></td>
</tr>
<tr>
<td></td>
<td>direct</td>
<td></td>
<td>inferred</td>
<td>reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 choices</td>
<td>visual</td>
<td></td>
<td>inferred</td>
<td>assumed</td>
<td>reported</td>
<td></td>
</tr>
</tbody>
</table>

Evidence type is a traditional and intuitive basis for proposals of the internal structure of evidential systems, but a truly universal typology of evidentiality must include the other dimensions of meaning that evidentials may encode, argues Matthewson (2011). She posits a typology based on three different dimensions of evidential meaning: evidence type, evidence location and evidence strength. The first dimension, evidence type, represents the by-now familiar set of parameters that classify evidence as direct, reported and so on. The second dimension, evidence location, encodes information about the location of the speaker in relation to the evidence. Specifically, this dimension distinguishes between whether the speaker witnessed the event itself or the results of the event. Evidence strength is the third dimension in this typology, and it expresses the trustworthiness or reliability of the evidence from the viewpoint of the speaker. Evidentials may encode direct and indirect values on any of the three dimensions, yielding a six-way typology (Table 2.5).
The typology predicts that evidential contributions may encode information about one, two or three of the dimensions. For example, consider the Tibetan evidential morpheme ‘dug. According to Matthewson’s system, ‘dug is a direct evidential whose meaning encodes directness in the two dimensions of evidence-type (sensory) and evidence-location (witness of the event itself rather than its results), but does not encode evidence-strength. Camacho (2012b) compares ‘dug’s evidential meaning to evidential readings of estar predicates in Spanish. Continuing the analogy, Matthewson’s typology makes the prediction that estar encodes directness in the two dimensions of evidence-type and evidence-location.

The theoretical modeling of the inventory of evidential types has also involved implicational hierarchies and scales, albeit less frequently than the categorical typologies discussed so far. Implicational hierarchies specify the set of evidential types of a language based on the presence or absence of other elements in the system. For instance, implicational hierarchies have been applied to classify types of sensory evidence across languages, with fairly uniform conclusions across studies (Aikhenvald, 2004; Speas, 2004; Willett, 1988). If a language specifies only one type of sensory evidence, then it is visual evidence, and if two types are specified, then these two are visual and auditory evidence. No language distinguishes among evidence perceived through taste, smell, and touch, but some languages encode a three-way contrast between visual, auditory and other sensory evidence.

Finally, the notion of degree of strength forms the basis for evidential scales (e.g., de Haan, 1998; Faller, 2002, 2012). Faller (2012) proposes a scale that orders evidential types according to implicatures of evidential strength. For example, a direct evidential contains a greater degree of illocutionary strength than a reportative evidential, which carries an implicature that the speaker lacks direct evidence for the proposition. This analysis is compatible with the third dimension of

Table 2.5. Three dimensions of evidential meaning (Matthewson, 2011, p. 4)

<table>
<thead>
<tr>
<th>Evidence type</th>
<th>Direct</th>
<th>Indirect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence location</td>
<td>event itself</td>
<td>results</td>
</tr>
<tr>
<td>Evidence strength</td>
<td>best</td>
<td>not best</td>
</tr>
</tbody>
</table>

| Evidence type       | e.g. sensory| e.g. inference, report |

- Evidence type: e.g. sensory; e.g. inference, report
- Evidence location: event itself; results
- Evidence strength: best; not best

The typology predicts that evidential contributions may encode information about one, two or three of the dimensions. For example, consider the Tibetan evidential morpheme ‘dug. According to Matthewson’s system, ‘dug is a direct evidential whose meaning encodes directness in the two dimensions of evidence-type (sensory) and evidence-location (witness of the event itself rather than its results), but does not encode evidence-strength. Camacho (2012b) compares ‘dug’s evidential meaning to evidential readings of estar predicates in Spanish. Continuing the analogy, Matthewson’s typology makes the prediction that estar encodes directness in the two dimensions of evidence-type and evidence-location.

The theoretical modeling of the inventory of evidential types has also involved implicational hierarchies and scales, albeit less frequently than the categorical typologies discussed so far. Implicational hierarchies specify the set of evidential types of a language based on the presence or absence of other elements in the system. For instance, implicational hierarchies have been applied to classify types of sensory evidence across languages, with fairly uniform conclusions across studies (Aikhenvald, 2004; Speas, 2004; Willett, 1988). If a language specifies only one type of sensory evidence, then it is visual evidence, and if two types are specified, then these two are visual and auditory evidence. No language distinguishes among evidence perceived through taste, smell, and touch, but some languages encode a three-way contrast between visual, auditory and other sensory evidence.

Finally, the notion of degree of strength forms the basis for evidential scales (e.g., de Haan, 1998; Faller, 2002, 2012). Faller (2012) proposes a scale that orders evidential types according to implicatures of evidential strength. For example, a direct evidential contains a greater degree of illocutionary strength than a reportative evidential, which carries an implicature that the speaker lacks direct evidence for the proposition. This analysis is compatible with the third dimension of
Matthewson’s (2011) typology, evidence strength, which classifies direct evidence as stronger than indirect. In Faller’s proposal, the evidential scale informs but one of the sincerity conditions at the illocutionary level of the sentence. As such, the proposal both suggests an internal structure for evidential systems and provides a formal mechanism for associating evidentials with the wider domain of illocutionary meaning.

2.6.4 Evidentiality and related meanings

The nature of the connections among the related semantic categories of evidentials, epistemics, evaluatives and speech acts is an open question. The reliability of the information source, the strength of the assertion, and the speaker’s judgments about the proposition itself are all associated with evidential meaning. Cross-linguistically, these meanings may be expressed by the same lexical or grammatical elements that encode evidentiality. Inferential evidentials present a particular challenge in this regard, as they involve both direct and indirect sources of evidence. Faller (2002), for instance, considers the category of inferred evidence to be the site of overlap between evidential and epistemic meaning in Cuzco Quechua. This contrasts with the categories of direct evidence and reported evidence, which fall squarely into the domains of direct and indirect evidence, respectively. The tremendous cross-linguistic variation in the encoding of evidentiality and epistemic modality has often hampered efforts to establish universals, but has also led to much fruitful debate.

In addition to epistemic modality, other categories of meaning related to evidentiality are evaluatives and presupposed knowledge forms. Evaluatives mark the viewpoint of the speaker. They include the category of mirativity, which is variously defined as the marking of information that is new to the speaker, the encoding of information about the speaker’s expectations concerning the proposition, and the expression of surprise. General knowledge forms may signify different contrasts with other forms in evidential systems. They can be associated either with assumption or visual evidence, depending on the system (Aikhenvald, 2004, p. 368). In

6 Faller’s terms for the Cuzco Quechua evidential markers are Direct, Conjectural and Reportative. Conjectural is the term that has traditionally been used to refer to inferred evidence in studies of Cuzco Quechua.
their work on evidentials in St’át’imcets, Matthewson, Davis, and Rullman (2007) identify two markers that express the meanings *contrary to expectations* and *presupposed knowledge*. The authors suggest that these morphemes are evaluatives, but that they might form part of the evidential system of St’át’imcets.

### 2.6.5 Semantics of evidentials

While much discussion has revolved around teasing apart the conceptual distinctions among evidentials, epistemics, evaluatives and speech acts, some studies have attempted to demonstrate the level of meaning an element contributes to using formal diagnostics (see de Haan, 1998, 2001; Faller, 2002, 2006; Matthewson et al., 2007; among others). The tests evaluate whether an evidential is associated with the illocutionary or propositional level of meaning. If an evidential contributes to the truth conditions of a sentence, it forms part of the propositional content, whereas an evidential that is associated with the illocutionary level of meaning functions as a comment by the speaker on the propositional content.

In addition to formal treatments of evidentials as either illocutionary modifiers or propositional operators, other types of non-propositional meaning, including presuppositions, conversational implicatures and conventional implicatures, are incorporated into some proposals or suggested as possible lines of inquiry. Faller (2002) advances an illocutionary-level analysis, in which evidentials in Cuzco Quechua modify the sincerity conditions of a speech act. In the propositional-level treatment of evidentials, Kratzer’s (1977, 1981) analysis of epistemic modals involving quantification over possible worlds has been influential. The possible-worlds analysis appears in a variety of studies of different languages, such as Garrett (2001) on Tibetan, McCready and Ogata (2007) on Japanese, and Matthewson et al. (2007) on St’át’imcets, to name a few. Izvorski (1997) analyzes evidentials in Bulgarian as epistemic modals using the possible-worlds approach, but with an additional meaning component in which evidentials carry a presupposition that there is evidence of a certain type for the proposition they embed. In a proposal that could be applied to analyze evidential meaning, Potts (2005) uses formal semantic tools to derive the meanings associated with conventional implicatures through semantic
composition. Recently, the presuppositional and conversational implicature analyses of evidentials have come under significant criticism (Speas, 2008). If other types of non-propositional meaning cannot account for evidential meaning, then a theory of conventional implicatures represents a promising line of inquiry for evidentials.

2.6.6 Summary: Evidentiality

In this section, I have reviewed some of the central concerns and findings in the literature on evidentiality. Establishing universals in this field has been particularly challenging due to the heterogeneity of evidential morphosyntax and semantics across languages. Typological studies of evidentiality (e.g., Aikhenvald, 2004; Chafe & Nichols, 1986; among others) have formed the basis for insights into the inventory of evidential types and the relations among the categories of evidentiality, epistemics, evaluatives and speech acts. However, these contributions have been criticized for the inconsistency of the terminology and descriptions used across studies. Formal approaches (e.g., Faller, 2002; Matthewson et al., 2007; among others) take advantage of the diagnostic tools of formal semantics to evaluate cross-linguistic data. The use of consistent methodologies has led to advances in the understanding of the levels of meaning to which evidentials contribute, as well as the types of analysis that are appropriate for different kinds of evidential systems.

In this study, I adopt a definition of evidentials in the broad sense, both in terms of their morphosyntax and their semantics/pragmatics. Evidentiality does not necessarily take the form of a morphological paradigm whose expression is obligatory. Evidentials can be lexical, grammatical or grammaticalized forms whose realization is optional, as is the case with estar. I analyze the evidential contribution of estar as an optional meaning extension, or as an “evidentiality strategy”, to use Aikhenvald’s (2004) term, and adopt the recurring parameters of direct/inferred/reported as the tripartite division of evidence type. I also find Blain and Déchaine’s (2007) interpretation of the broad definition of evidentiality, which characterizes evidentials as markers of the viewpoint of the speaker, to be relevant in characterizing the evidential use of estar.
In this section, I have considered how the evidential use of *estar* can be situated within the larger context of the theoretical and typological literature on evidentiality. Previous studies on this topic have consisted of theoretical contributions, and the current study will add an empirical analysis to the pre-existing body of work on the evidential use of *estar*. In the next chapter, I consider previous empirical studies of the use of *ser* and *estar* with adjectives, and review the findings of research on variation and change in this context.
Chapter 3
Variation and change in the use of ser and estar with adjectives

3.0 Introduction

In the previous chapter, I considered a range of analyses which covered different dimensions of the distribution and interpretation of ser and estar with adjectives, including the various elements in the structure, the diverse types of clauses that they comprise, and a host of pragmatic factors that all may contribute to the meaning of a given sentence. This complex picture is complemented by a well-documented history of variation and change in the use of ser and estar with adjectives. The phenomenon continues to evolve in the present day, and data from recent sociolinguistic studies has much to offer to our understanding of the grammar. In this chapter, I review the findings of the main recent sociolinguistic studies of ser and estar, all of which investigate some combination of linguistic and extralinguistic variables in ser/estar selection in one or more varieties of Spanish.

3.1 Diachronic studies

Several Romance languages have a dual copula system, whose historical origin can be attributed to the Latin verbs essere/sedere and stare (Spanish ser and estar, respectively). All of the Iberian Romance languages (Catalan, Galician, Spanish and Portuguese) share a similar path of historical development characterized by the extension of stare into contexts that originally featured essere. In Spanish, the extension of estar into contexts previously occupied by ser is not a recent phenomenon (see for example Hengeveld, 1992; Penny, 1991; Pountain, 1982).
Vaño-Cerdá’s (1982) diachronic study of the copulas in adjectival contexts dates the earliest cases of the extension of *estar* to the 12th century. Vaño-Cerdá describes four major stages in the extension of *estar*. In the first stage, around the 12th to 13th century, *estar* appears with adjectives that express non-inherent spatial, physical and psychological characteristics of the subject (79). The second and third stages, from the 15th century onwards, are characterized by the occurrence of *estar* with adjectives expressing a change of state (80) and the use of *estar* with adjectives that describe the current condition or behaviour of the subject (81). Speaker reactions to an immediate experience with *estar* are noted in the fourth stage (82), which begins in the 16th to 17th century.

(79)  
Stage I: Non-inherent spatial, physical and psychological characteristics of the subject

¿Que as que estas triste e cuydosa?

‘What have you that you are ESTAR sad and distressed?’

   (Calila, cited in Vaño-Cerdá, 1982, p. 238)

(80)  
Stage II: Change of state

Por donde conoció que su vecino (= d. Quijote) estaba loco.

‘Where he learned that his neighbor (don Quijote) was ESTAR crazy.’

   (Cervantes, cited in Vaño-Cerdá, 1982, p. 281)

(81)  
Stage III: Current condition or behaviour of the subject

(Tal como habla, tal como se comporta) ¡Jesu, Jesu! ¡Qué fiero que está!

‘(As one speaks, so one behaves) Jesus, Jesus! How fierce he/she is ESTAR!’

   (Celestina, cited in Vaño-Cerdá, 1982, p. 307)

(82)  
Stage IV: Speaker reactions to an immediate experience

Señor, este pan (que me estoy comiendo) está sabrosísimo... ...; pruebe,
señor, y verá qué tal está.

‘Sir, this bread (that I am eating) is very tasty … taste it, sir, and you shall see that it so.’

(Lazarillo, cited in Vaño-Cerdá, 1982, p. 286)

Of particular relevance to this study is the sentence in (82), which I classify as an evidential use of *estar*. Sentences (79), (80) and (81) could give rise to evidential readings as well. However, they have animate subjects, and without more information about the surrounding context, the evidential reading is not salient. More recently, variability in contexts such as those in (79-82) has continued to form the basis for sociolinguistic studies, the focus of which has been the documentation and explanation of the extension of *estar*.

### 3.2 Recent studies: Methodologies

#### 3.2.1 Populations

Recent empirical studies have examined variation and change in the use of *ser* and *estar* in both contact and non-contact varieties. Investigations of Spanish-English bilinguals have contributed valuable data on this topic. Silva-Corvalán’s (1986) pioneering work on Mexican-American bilinguals in Los Angeles sparked a wave of studies on this topic from a sociolinguistic perspective. Since Silva-Corvalán, several other investigations have looked at copula choice in Spanish-English bilingual populations (Brown & Cortés-Torres, 2012; Ortiz-López, 2000; Salazar, 2007). In addition to Spanish-English bilingual populations, several studies have been conducted with monolingual Spanish populations (Alfaraz, 2012; Cortés-Torres, 2004; de Jonge, 1993; Díaz-Campos & Geeslin, 2011; Gutiérrez, 1992). Sociolinguistic studies of *ser* and *estar* have also been carried out with bilingual speakers of Spanish and a contact language other than English (Aguilar-Sánchez, 2012; Geeslin & Guijarro-Fuentes, 2007, 2008). Table 3.1
summarizes the varieties and contact languages, as well as the instruments used for data collection, of the studies reviewed here.

Table 3.1. Previous studies: Populations and data collection methods

<table>
<thead>
<tr>
<th>Study</th>
<th>Variety and contact languages</th>
<th>Data collection method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monolingual speakers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gutiérrez (1992)</td>
<td>Morelia, Mexico</td>
<td>Interview</td>
</tr>
<tr>
<td>de Jonge (1993)</td>
<td>Mexico and Venezuela</td>
<td>Corpus</td>
</tr>
<tr>
<td>Cortés-Torres (2004)</td>
<td>Cuernavaca, Mexico</td>
<td>Questionnaire and interview</td>
</tr>
<tr>
<td>Díaz-Campos and Geeslin (2011)</td>
<td>Venezuela</td>
<td>Corpus</td>
</tr>
<tr>
<td>Alfaraz (2012)</td>
<td>Cuba</td>
<td>Corpus</td>
</tr>
<tr>
<td><strong>Spanish-English bilinguals</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silva-Corvalán (1986)</td>
<td>Los Angeles, U.S.</td>
<td>Interview</td>
</tr>
<tr>
<td><strong>Bilinguals in Spanish and a language other than English</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geeslin and Guijarro-Fuentes (2007, 2008)</td>
<td>Spain</td>
<td>Contextualized preference task</td>
</tr>
<tr>
<td></td>
<td>Contact languages: Galician, Valencian, Catalan, and Basque</td>
<td></td>
</tr>
<tr>
<td>Aguilar-Sánchez (2012)</td>
<td>Limón, Costa Rica</td>
<td>Interview</td>
</tr>
<tr>
<td></td>
<td>Contact languages: English, Cantonese, Bribri, French and Arabic</td>
<td></td>
</tr>
</tbody>
</table>

3.2.2 Quantitative approaches
In general, two major trends in quantitative methodology can be identified across the studies. Earlier projects measured the frequency of the innovative use of *estar*, and binary quantitative analyses contrasted rates of innovative versus prescriptive uses of this copula (Cortés-Torres, 2004; Gutiérrez, 1992; Ortiz-López, 2000; Silva-Corvalán, 1986). Later studies have adopted a multivariate analysis, where findings are presented as rankings of predictive factors of copula choice in terms of significance and strength (Aguilar-Sánchez, 2012; Alfaraz, 2012; Brown & Cortés-Torres, 2012; Díaz-Campos & Geeslin, 2011; Geeslin & Guijarro-Fuentes, 2007, 2008; Salazar, 2007). This methodological shift is likely a consequence of changing approaches to quantitative analysis in the broader field of sociolinguistics. Another approach has employed a usage-based model, where lexical items are grouped into semantic clusters in a functional analysis of the [ser/estar + adjective] context. I consider the advantages and limitations of each approach.

3.2.2.1 Binary analysis: Comparing prescriptive and innovative uses of *estar*

The studies that quantified rates of production or selection of the innovative use of *estar* have adopted various definitions of what qualified as an innovative context with this verb. For instance, Silva-Corvalán (1986) mentions two criteria that guided her coding decisions of innovative versus conservative uses of *estar*. The first is Falk’s (1979) class versus individual norm (or frame of reference), in which *ser* is associated with the former, and *estar*, the latter (see section 2.3.3 for discussion). The second is the concept of “susceptibility to change”. This variable is not defined, but the reader is referred to Falk (1979) and Navas-Ruíz (1977) for details. Silva-Corvalán presents an example of an innovative use of *estar* (83).


*My father was a very tall man. “All the Campas are tall” – as my...*
uncle once told me – “except you, Daniel.” (I: He really told you so?) The son of a bitch! And I told him: “But I’m ESTAR intelligent and very handsome and I can’t have everything.””

(1986, p. 593)

She explains her procedure for coding the innovative versus conservative uses of estar: “The 344 cases of innovative use of estar were so classified on the well-tested assumption that the speaker did not intend to convey any meaning difference by choosing estar rather than ser” (Silva-Corvalán, 1986, p. 597).

Gutiérrez (1992) classified estar for innovative versus prescriptive uses, and the prescriptive use of estar is described as a context “where the copula is expressing the state of a thing at a certain moment” (1992, p. 116). Although he does not explicitly define an innovative use of estar, he presents the following as an example (84).

(84) ¿Cómo crees tú que es los Estados Unidos?

‘What do you think of the U.S.?’

Pues dicen que hay, que hay muchos árboles frutales, que hay muchos vegetales allá y ... no comen grasas, que están muy ... que están muy altos allá ...

‘It is said that there are HABER, that there are HABER a lot of fruit trees, that there are HABER a lot of vegetables there and ... they do not eat fat, that they are ESTAR very ... that they are ESTAR very tall there...’

(Gutiérrez, 1992, p. 116)

He goes on to note that the contrast between individual and class norm is relevant to distinguishing innovative from prescriptive uses of estar. According to prescriptive usage, ser occurs in a class norm context and estar in an individual norm context. When estar appears in a context that expresses a class norm, the sentence can be considered an innovative use (85).
(85) ... como el camión era muy grande tenía asiento de sobra ... no, ... como el carro de mi hijo también está muy grande, ... tiene metro y medio más grande que todos ...

‘... since the bus was very big it had lots of seats ... no, ... like my son’s car that is also very big, ... it’s a meter and a half longer than the others ...’

(Gutiérrez, 1992, p. 118)

De Jonge (1993) classified all cases of estar in combination with an age adjective as innovative (86).

(86) Juan está joven/viejo.

‘Juan is young/old.’

(de Jonge, 1993, p. 71)

Ortiz-López (2000) provides examples of the extension of estar in traditional contexts of ser, such as those in (87).

(87) a. Si se van unos pocos (de Puerto Rico), la comida va a estar más abundante.

‘If just a few leave (Puerto Rico), food will be more plentiful.’

b. Esa nariz está bien fea.

‘That nose is pretty ugly.’

c. Después que decidí salir con la chica, la tipa estaba fea, la tipa estaba bien fea.

‘After I decided to go out with the girl, the girl was ugly, the girl was pretty ugly.’
d. *Los aguaceros hoy no estarán tan numerosos.*

‘Downpours aren’t *estar* as common today.’

(Ortiz-López, 2000, p. 100)

Alfaraz (2012) adopted a definition of an innovative use of *estar* in which innovation is defined in terms of semantic bleaching. She argues that a sentence like *Julio está viejo* (‘Julio *is* old’) should only be classified as innovative if it expresses the same meaning as its minimal pair counterpart with *ser*, *Julio es viejo* (‘Julio *is* old’). She presents an example of a token (88) that is excluded from being classified as innovative on the basis that it expresses an individual frame of reference: “… the comparison is with the woman herself at different points in time” (Alfaraz, 2012, p. 10).

(88) *Y yo le dije no yo estoy muy fea ... Digo no, no tengo muchas canas y estoy muy fea.*

‘I told him no I am *estoy* too ugly … I said no, no I have too much gray hair and I am *estoy* too ugly.’

(Alfaraz, 2012, p. 10)

There are several limitations to the methodological approaches of these studies. The main problem concerns the subjective assessment of tokens. Even though the studies ensured inter-rater reliability, the variation and change affecting this structure mean that the definition of innovation is something of a “moving target”. Indeed, Díaz-Campos and Geeslin (2011) note that an approach that relies on coding for innovative versus prescriptive uses of *estar* suffers from the troublesome prospect of determining what constitutes prescriptive versus innovative usage at any given moment: “… it [the copula contrast] has been undergoing change for such a long period of time and, consequently, it is difficult to document a given prescriptive norm at any single point in time” (2011, p. 74). In addition, when criteria are made explicit, they may be left undefined, as is the case for the variable *susceptibility to change* in Silva-Corvalán’s study. Several studies (e.g. Alfaraz, 2012; Gutiérrez, 1992; Silva-Corvalán, 1986) converged on the notion of a class
norm with *estar* as an innovative context. However, given that additional criteria differed from one study to the next, it is possible for the same context to be classified as innovative in one study and prescriptive in another. If this is the case, quantitative findings are not directly comparable. Finally, different interpretations of *estar* predicates are indistinguishable in this type of analysis. In the sample data, the sentences in (87b) and (87c) may be evidential uses of *estar* rather than cases of semantic bleaching.

### 3.2.2.2 Multivariate analysis: Examining multiple linguistic and extralinguistic factors

Several studies mentioned here (Aguilar-Sánchez, 2012; Brown & Cortés-Torres, 2012; Díaz-Campos & Geeslin, 2011; Geeslin & Guijarro-Fuentes, 2007, 2008) adopted a methodology that included a multivariate analysis. Regression analyses were conducted using Goldvarb (Sankoff, Tagliamonte, & Smith, 2005) or Rbrul (Johnson, 2009) and involved various combinations of linguistic and extralinguistic variables.

A multivariate analysis differs from a comparison of prescriptive versus innovative uses of *estar* in examining the frequencies of copula selection in terms of different combinations of the categories of independent variables. It accommodates the fact that more than one factor may be responsible for the choice of *ser* or *estar* in any given context, and that there are almost always multiple variables at play. This type of analysis acknowledges that the various factors appear in different combinations, and it allows for examination of the interaction between linguistic factors and extralinguistic factors. A methodological approach that offers a nuanced look at variable contexts is superior to the subjective classification of tokens as either innovative or prescriptive as a tool for the study of copula selection in Spanish.

### 3.2.2.3 A usage-based approach
Brown and Cortés-Torres (2012) applied a usage-based approach, following Bybee and Torres Cacoullos (2008). In a usage-based approach, the basic unit of analysis is the construction, which in this case is \([\text{ser/estar} + \text{adjective}]\). Constructions have pragmatic meanings that are independent of the meanings of the individual lexical items they are composed of. Frequency plays an important role in the emergence of constructions: through repeated patterns of use, speakers come to associate form and meaning. These form-meaning correlations may become productive, as is the case with certain \([\text{estar} + \text{adjective}]\) contexts. The authors classified innovative and non-innovative uses of the copulas as follows. First, they established the canonical uses of each \([\text{ser/estar} + \text{adjective}]\) combination via a multivariate analysis using the linguistic factors frame of reference, adjective class, resultant state and experience with the referent. Then, innovative uses of \([\text{estar} + \text{adjective}]\) were defined as those that fall outside the canonical uses as determined by one or more of the linguistic predictors. The adjectives that appeared in these uses were organized into semantic clusters or clouds. A high-frequency lexical item was identified as a central member of a category (e.g. bueno, ‘good’), and other semantically related lexical items were added to the cluster (e.g. brutal, chévere, nitido, nice, tripioso, rico; Brown & Cortés-Torres, 2012, p. 69). In this example, evaluative adjectives with positive connotations were grouped together.

By combining a multivariate analysis with a usage-based model in their approach, Brown and Cortés-Torres have made a novel contribution to the study of the innovative use of estar. The authors used the findings of their multivariate analysis to identify semantic classes of adjectives that occurred in the innovative \([\text{estar} + \text{adjective}]\) construction. They tracked the lexical diffusion of the innovative use of estar with a usage-based approach, which has added a descriptive dimension to our understanding of what types of adjectives appear in innovative uses of estar in a given variety. However, there are limitations to a usage-based approach. It is hard to see how this approach could predict the direction of variation and change in the distribution of ser and estar with adjectives across dialects. In the absence of the recognition of syntactic and semantic constraints on the \([\text{ser/estar} + \text{adjective}]\) combination, frequency of use is the sole determinant of which adjectives enter into these contexts. The contribution of a usage-based approach is ultimately descriptive, whereas an analysis rooted in a theory of the formal syntactic and semantic properties of the elements and structures involved has predictive power.
Sociolinguistic studies of variation and change in the [ser/estar + adjective] context have employed various methodologies and approaches, including binary analysis of innovative versus prescriptive usage, multivariate analysis, and usage-based models of lexical diffusion. I now turn to the results of these studies. In the next section, I consider which linguistic variables are found to be significant predictors of the selection of estar, and whether any of them are consistently significant across the varieties studied.

3.3 Results of previous studies: Linguistic variables

3.3.1 Linguistic variables: Item-level variables

Item-level variables refer to individual lexical items that appear in the sentence, which in this context are adjectives. The most significant predictor of estar selection across all studies was adjective class, and indeed it is the only variable that appeared in all studies discussed here. Most employed some version of Dixon’s (1977, 2004) semantic classification of adjectives for purposes of categorization (see section 2.4.1 for discussion). The results of these studies revealed an overall pattern in the adjective classes that tended to appear in innovative contexts with estar, or in the case of multivariate analyses, that favoured the selection of estar.

Silva-Corvalán (1986) identified four classes of adjectives that favoured the innovative use of estar among Mexican-American bilinguals in Los Angeles. These included size adjectives such as alto (‘tall’) and grande (‘big’), adjectives depicting physical appearance with animate subjects (e.g., curioso ‘cute’), adjectives expressing age (e.g., joven ‘young’, mayor ‘old’), and adjectives of evaluation, like barato (‘cheap’), bueno (‘good’), difícil (‘difficult’), fácil (‘easy’), and interesante (‘interesting’). Other categories of adjectives occurred with estar in innovative contexts, but with lower frequencies. These were items that expressed “sensory character”, that is, adjectives related to the senses of taste, smell and hearing such as dulce (‘sweet’), caliente (‘hot’), and silencioso (‘quiet’), and adjectives that described the physical appearance of non-animate subjects (e.g., tosco ‘coarse’, liso ‘smooth’, sucio ‘dirty’).
Some of these findings can be interpreted with a new awareness of the evidential use of *estar*. The cases of adjectives that express the sensory character of an entity with *estar* are potentially evidential contexts, in that the senses of taste, smell, hearing, and so on are involved. Also, Silva-Corvalán observes that evaluative adjectives often appeared in contexts that matched descriptions of the evidential use of *estar*. These items occurred frequently in the evaluation of entities found in the immediate context, and in expressions of subjective reactions where *estar* is the accepted copula (e.g. ¡Qué difícil está esto! ‘How difficult this is!*ESTAR!*’, *Está buena esta naranja*, ‘This orange *is* *ESTAR* good.’) (Silva-Corvalán, 1986, p. 599). If the selection of *estar* with sensory and evaluative adjectives encodes the perspective of the speaker, then these contexts correspond to the definition of the evidential use of *estar*.

Like Silva-Corvalán, Gutiérrez found that the categories of adjectives that showed the highest rates of the innovative use of *estar* were age, size, physical appearance and evaluation in the monolingual Spanish of Morelia, Mexico. He further notes:

> It would appear that an element of subjectivity in these adjectives prompts the appearance of *estar*. The meaning of these adjectives is related to the opinions of the speaker regarding persons, animals or things according to parameters that he/she establishes to make the qualification. (Gutiérrez, 1992, p. 127)

Gutiérrez draws a parallel between the uses of innovative *estar* and the diminutive suffix –*ito* as discursive strategies that encode subjectivity, and presents an example where the two coincide (89).

(89)  ... otro lugar que podría ser bueno, el dueño de este lugar tiene otro en el centro

> que está más *pequeñito*, más *encerradito*, podría decirse que, tal vez, un poquito más íntimo, se llama ‘La Tórtola’, ahí hacen el mismo tipo de comida ... muy sabrosa ...

‘… another place that might be *SER* good, the owner of this place has another one in downtown that *is* *ESTAR* smaller*DIM*, cozier*DIM*, it could be said, perhaps, a little more intimate, it is called ‘La Tortola’, they make the same kind of food … very ...
Ortiz-López (2000) compared the frequencies of different adjective classes in the innovative use of *estar* among bilinguals in Puerto Rico using a written questionnaire and interviews. For the questionnaire, the results showed that innovative *estar* appeared most frequently with adjectives expressing state, age, size and evaluation, and in the interviews, with adjectives of age and size. In Cortés-Torres’s (2004) data, the extension of *estar* occurred most frequently with adjectives of physical appearance, age, size and evaluation in the monolingual Spanish of Cuernavaca, Mexico. Salazar (2007) identified the adjective classes of evaluation, physical description/states of being, descriptions of colour and size for inanimate subjects and age as those that favoured *estar* selection in the northern New Mexico/southern Colorado variety of Spanish. Díaz-Campos and Geeslin’s (2011) results indicated that adjectives expressing mental states, physical states and social status (e.g., *casado* ‘married’) strongly favoured *estar* in monolingual Caracas Spanish.

In their study of monolingual and bilingual speakers in four regions of Spain, Geeslin and Guijarro-Fuentes (2008) found that six items in the contextualized preference task elicited nearly unanimous *estar*-responses across groups or the unanimous selection of *estar* in at least one group. Four of the six contexts involved adjectives that refer to mental states: *enojado* (‘mad’), *acostumbrado* (‘used to’), *alegre* (‘happy’), and *despierto* (‘aware’). The two other contexts contained an adjective denoting a physical state, *enfermo* (‘sick’), and an adjective denoting a sensory characteristic, *frío* (‘cold’).

Adjectives that refer to physical properties and value (e.g., *bueno* ‘good’, *bello* ‘beautiful’, *cómodo* ‘funny’) favoured *estar* in Alfaraz’s (2012) analysis of the Cuban variety. Some contexts where *estar* appeared with these categories of adjectives could give rise to evidential readings, such as those in (90).

(90) a.  
   *El barco estaba fuerte.*

   ‘The boat was **ESTAR** strong/sturdy.’
b. *Estaban bonitos los cuadros que hizo la señora de ( ) ... los de los pájaros estaban lindos, me gustaron.*

‘The paintings that the woman from ( ) made were pretty … the ones with the birds were pretty, I liked them.’

c. [La mujer] no está fea. Tiene facciones bonitas.

‘[The woman] is not ugly. She has pretty features.’

d. *La película está simpática.*

‘The movie is amusing.’

e. *La película está cómica.*

‘The movie is funny.’

(Alfaraz, 2012, p. 20)

An evidential reading is particularly salient in the use of *estar* with an evaluative (value) adjective in (91).

(91) *Ay los árboles son en esa época ... nosotros fuimos en octubre y los árboles – hay veces que usted veía un árbol al lado del otro y no tienen – dos hojas no tienen el mismo color ... un árbol es rojo y el otro que tiene al lado es amarillito y el otro verde. *Está* divino.*

*Divino, divino.*

‘Oh the trees during that season … we went in October and the trees – sometimes you would see one tree next to another and they don’t have – two leaves don’t have the same colour … one tree is red and the other next to it is yellow and the other green. *It*
is ESTAR divine. Divine, divine.’

(Alfaraz, 2012, p. 19)

The selection of estar here expresses the perspective of the speaker in recalling and evaluating a firsthand experience.

Brown and Cortés-Torres (2012) showed that estar was favoured with adjectives denoting mental and physical states and disfavoured with adjectives of evaluation, age and size. They determined that the most frequent adjectives to appear in the [estar + adjective] construction were brutal (‘brutal’, with both positive and negative senses), bueno (‘good’), bonito (‘pretty’), malo (‘bad’), fuerte (‘strong’), chévere (‘cool’) and caro (‘expensive’). High-frequency adjectives such as these acted as central members of semantic clusters of related adjectives, which they suggest as the basis for the productive use of the construction through analogy. Four semantic clusters, those based on the central-member adjectives brutal, bueno, bonito and chiquito (‘small’), accounted for 95% of the innovative uses of estar in their study. It is worth noting that all of these are gradable adjectives, and many of the contexts in which they appeared are ambiguous between semantic bleaching between the copulas (referred to as innovation) and the evidential use of estar.

Other item-level variables that appeared in these studies included circumstantiality, resultant state and “copulas allowed”. Silva-Corvalán (1986) defines circumstantiality as a distinction between an inherent or essential property as compared with an accidental or circumstantial property expressed by an adjective. This variable was significant in predicting the innovative use of estar in her data. Resultant state adjectives (Bosque, 1990) were significant predictors of estar selection in the studies by Díaz-Campos and Geeslin (2011) and Brown and Cortés-Torres (2012), but not in Aguilar-Sánchez’s (2012) study. Geeslin and Guijarro-Fuentes (2008) employed the copulas allowed variable to separate categorical from variable [ser/estar + adjective] contexts. This factor sorted adjectives into three classes according to the verb(s) they combined with: ser only, estar only, and both allowed. Copulas allowed was significant in predicting the choice of estar.
In summary, the various item-level factors examined in these studies included four significant predictors of estar selection: adjective class, circumstantiality, resultant state and copulas allowed. The most salient finding across varieties was that the semantic class of an adjective is a powerful determinant of its propensity to combine with estar, specifically in the case of adjectives of age, size, physical appearance and evaluation. Several of these previous studies have identified as innovative cases in which evaluative adjectives (e.g. bueno ‘good’, bonito ‘pretty’) appeared with estar, but which I analyze as instances of the evidential use of estar. Nonetheless, given that the process of categorization of adjectives into semantic classes relied on subjective judgments by researchers rather than formal diagnostics, there was no agreement across studies in the definitions and values assigned to the categories. There is also the problem of indeterminacy, where adjectival meaning may be underspecified without additional context (Kennedy, 1999). For example, in (92) the adjective grande (‘big’) could be classified either as an adjective of population size or physical property, depending on the dimension being measured.

(92)  
La ciudad es grande.

‘The city is SER big.’

These inconsistencies in the classification of adjectives make it difficult to compare quantitative results across studies.

3.3.2 Linguistic variables: Sentence-level variables

At the level of the sentence, the studies examined whether the following variables favoured the selection of estar: predicate type, animacy of the subject, intensifiers, temporal adverbs, norm/frame of reference, and susceptibility to change. Geeslin and Guijarro-Fuentes (2008) and Díaz-Campos and Geeslin (2011) adopted a theory of the individual-level versus stage-level predicate distinction that defined the contrast according to the temporal extension of the predicate (Carlson, 1989; Fernández-Leborans, 1999; Leonetti, 1994). These authors found that predicate type was a strong predictor of estar, whereas Aguilar-Sánchez (2012) did not. Silva-
Corvalán (1986) and Cortés-Torres (2004) looked for evidence that animate subjects favoured the innovative use of *estar*, but their data did not support this prediction. Salazar’s (2007) analysis showed that intensifiers such as *muy* (‘very’) and *más* (‘much’) favoured the choice of *estar*.

In his analysis of *estar* with age adjectives, de Jonge (1993) identified the temporal adverbs *ya* (‘already’) and *cuando* (‘when’) as co-occurring elements in contexts of the semantic extension of *estar*. Ortiz-López (2000) observed that temporal adverbs such as *ahora* (‘now’), *cuando* and *más que antes* (‘more than before’) appeared in contexts of the extension of *estar* with certain semantic classes of adjectives: age, size, physical appearance, class and evaluation. Two multivariate analyses (Aguilar-Sánchez, 2012; Salazar, 2007) measured the presence of a temporal adverb as a conditioning factor in the selection of *estar*. The results of both studies confirmed that a temporal adverb in the context was a significant predictor of *estar*. These findings lend empirical support to formal proposals regarding the temporal use of *estar* with adjectives (see section 2.3 for discussion), a meaning which may overlap with the evidential use of *estar* in certain contexts.

Another sentence-level factor that these studies investigated was whether a *ser/estar* context expresses an individual or class norm/frame of reference (Clements, 1988; Falk, 1979). Four studies that measured norm/frame of reference found that it was a significant predictor of *estar* in the Los Angeles, Morelia, Caracas and Puerto Rican varieties (Brown & Cortés-Torres, 2012; Diaz-Campos & Geeslin, 2011; Gutiérrez, 1992; Silva-Corvalán, 1986). Geeslin and Guijarro-Fuentes’s (2008) data showed that norm/frame of reference was not a significant predictor of *estar* in the Iberian varieties tested in their study.

For Silva-Corvalán and Gutiérrez, frame of reference was viewed as the root cause of the meaning difference between *ser* and *estar* contexts, and it was the main criterion for distinguishing an innovative use of *estar*. When *estar* appeared in a class frame of reference, it was classified as an innovative use (93).

(93)  ... *como el camión era muy grande tenía asiento de sobra... no,... como el carro de mi hijo también está muy grande,*... tiene metro y medio más grande que todos...
‘... since the bus was \textsc{ser} very big it had lots of seats... no,... like \textbf{my son’s}
\textbf{car that is} \textsc{estar also} \textbf{very big}... it’s a meter and a half longer than the others...’

(Gutiérrez, 1992, p. 118)

The categorization of evidential uses of \textit{estar} in terms of norm/frame of reference is problematic.
According to the norm/frame of reference variable, \textit{[estar + adjective]} contexts that express the viewpoint of the speaker, such as evaluative adjectives in evidential uses of \textit{estar} (2), are classified as innovative uses of \textit{estar} on the basis that they do not express a comparison between the subject and itself.

(2) a. \textit{Este vino está riquísimo.}

‘This wine \textsc{is} \textit{estar} delicious\textit{INT}.’

b. \textit{¿Viste la nueva película de Almodóvar? Está buena.}

‘Did you see the new Almodovar film? It \textsc{is} \textit{estar} good.’

c. \textit{Están muy lindos los cachorros del vecino.}

‘The neighbor’s puppies \textsc{are} \textit{estar} cute.’

Such a classification system would include these uses with cases of semantic bleaching between the copulas, and would miss out on identifying the evidential meaning conveyed by the choice of \textit{estar} over \textit{ser} in these contexts.

Furthermore, the class versus individual categories of comparison may not capture the whole range of comparative possibilities. Franco and Steinmetz (1983, 1986) also described the \textit{ser/estar} distinction in terms of the notion of comparison, but their analysis presented a different matrix of possibilities, including a comparison between a speaker’s prior expectations and their firsthand experience. This type of comparison (or contrast between expectations and experience) corresponds to the definition of mirativity, which is closely related to evidentiality (see section 2.6.4 for discussion). Franco and Steinmetz’s approach to the different types of comparison
expressed by [ser/estar + adjective] contexts extends to evidential uses of estar in a way that the norm/frame of reference analysis cannot.

The findings for the factor susceptibility to change varied across studies. This variable characterizes the relationship between the subject and the adjective in a ser/estar clause, and its [+/-] value reflects whether this relationship encodes a capacity for change. Silva-Corvalán (1986) determined that it was a key criterion for the loss of the semantic contrast between ser and estar observed among Mexican-American bilinguals in Los Angeles. Additionally, Geeslin and Guijarro-Fuentes (2008) found that it was a significant predictor of estar for the monolingual group and two out of four bilingual groups of Peninsular Spanish speakers in their study. Díaz-Campos and Geeslin’s (2011) data showed that estar was favoured with “changeable” adjectives in monolingual Caracas Spanish. However, two other studies found that estar was not favoured with changeable adjectives in Puerto Rican and Costa Rican varieties (Aguilar-Sánchez, 2012; Brown & Cortés-Torres, 2012).

An updated formal treatment of many of the variables included in these previous studies is available. The variables of norm/frame of reference and susceptibility to change can be reframed in terms of the components of gradability as presented in semantic compositional approaches to the quantization of states (Gumiel-Molina & Pérez-Jiménez, 2012; Husband, 2012; Kennedy & McNally, 2005). A gradability analysis considers the composition between an adjective’s scale structure and the standard of comparison expressed in the context (see section 2.4.3 for discussion). This approach requires a revision of what previous studies have referred to as susceptibility to change and norm/frame of reference in terms of the two variables scale structure and standard of comparison. Susceptibility to change is a way of describing the mapping of an adjective’s scale structure onto the subject, and its value is either open-scale or closed-scale. The notion of comparison encoded by the norm/frame of reference variable is represented in the standard of comparison present in the semantic composition, and it is valued as absolute or relative. Furthermore, the variable presence of an intensifier is also captured by a gradability analysis. Intensifiers (e.g. muy ‘very’, más ‘more’), also known as degree modifiers and comparatives, are grammatical with gradable adjectives (94a), but not with non-gradable adjectives (94b).
This approach to reframing these variables in terms of recent theories about the quantization of stative predicates presents an opportunity to test the predictions of proposals about the distribution of gradable adjectives against data for *ser/estar* distribution with adjectives in Spanish.

### 3.3.3 Linguistic variables: Discourse-level variables

Discourse-level variables encompass the pragmatic aspects of copula selection with adjectives, including elements of the utterance context such as speaker, hearer and deixis. The discourse-level variables in the studies under consideration here include semantic transparency and experience with the referent. These pragmatic factors were found to be significant predictors of the selection of *estar*, and are highly relevant to an analysis of the evidential use of *estar*.

Silva-Corvalán (1986) and Ortiz-López (2000) observed that semantic transparency was significantly correlated with the innovative use of *estar*. Silva-Corvalán explains the origin of the concept of semantic transparency:

… the process [of gradual diffusion] appears to be controlled by what I have called the ‘semantic transparency’ of the choice between *ser* and *estar* to introduce an adjective in a given sentence. The motivation for including this factor is provided by the observation, made in historical linguistics, that complex systems are more likely to change than simple ones; i.e. when the rules underlying an aspect of the grammar are opaque, we have a potential source for reinterpretation. (1986, p. 600)
The values of the semantic transparency variable and the corresponding frequency of occurrence of the innovative use of estar were different modality, apparent synonymy, choice not allowed and clear difference (95).

(95)  a. Different modality

*El mole poblano está/es bueno ahí.*

‘Puebla mole (a Mexican dish) is*estar/ser good there.’

b. Apparent synonymy

*Si el hombre está/es soltero, puede hacer lo que quiera.*

‘If the man *estar/ser unmarried, he can do whatever he pleases.’

c. Clear difference

i. *Pedro es callado/vivo.*

‘Pedro is*ser shy/sharp.’

ii. *Pedro está callado/vivo.*

‘Pedro is*estar silent/alive.’

d. Choice not allowed

*Eso es/*está preferible.*

‘That is*ser/*estar preferable.’

(Silva-Corvalán, 1986, p. 601)

Based on this description, we can observe that semantic transparency is not a single variable, but rather is made up of a number of variables. Let us consider each of the categories of semantic transparency in order to determine which individual variable or set of variables they might encode. Choice not allowed is similar to the copulas allowed variable (Geeslin & Guijarro-
Fuentes, 2008), in that both variables attempt to formalize the distinction between categorical and variable contexts of *ser/estar* selection. The choice not allowed designation represents contexts where the choice is categorical, such as sentences where *ser* appears in combination with an epistemic adjective. Similarly, the clear difference classification corresponds to cases of polysemy of adjectives, where again *ser/estar* selection is categorical and a different lexical entry for a given adjective combines with either *ser* or *estar* to produce a specific meaning. Contexts of apparent synonymy are variable but show no contrast in meaning between the use of one verb or the other. Adjectives of status (e.g. *soltero* ‘single’, *casado* ‘married’) are an example of this semantic bleaching. The significance of these factors highlights the importance of distinguishing categorical from variable contexts of *ser/estar* selection.

Silva-Corvalán’s category different modality included, but did not quantitatively distinguish, many syntactic, semantic and pragmatic contrasts expressed by the choice of *ser* or *estar* in combination with an adjective. The set of parameters that she cites to account for the variability between *ser* and *estar* appear in Table 3.2.

Table 3.2: Parameters for copula selection (Silva-Corvalán, 1986, p. 590)

<table>
<thead>
<tr>
<th></th>
<th><em>ser</em></th>
<th><em>estar</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>inherent or essential</td>
<td>accidental or circumstantial</td>
<td></td>
</tr>
<tr>
<td>imperfective</td>
<td>perfective</td>
<td></td>
</tr>
<tr>
<td>permanent</td>
<td>temporary</td>
<td></td>
</tr>
<tr>
<td>defining, abstract and independent of immediate experience</td>
<td>dependent on concrete and/or immediate experience</td>
<td></td>
</tr>
<tr>
<td>not susceptible to change</td>
<td>susceptible to change</td>
<td></td>
</tr>
<tr>
<td>presented within a class frame of reference</td>
<td>presented within an individual frame of reference</td>
<td></td>
</tr>
</tbody>
</table>

One of the variable contexts that falls under the different modality heading is relevant to an inquiry into the evidential use of *estar*. The example sentence for different modality in (95a)
involves an evaluative adjective and a meaning difference that hinges upon the use of \textit{estar} to emphasize the experience of the speaker (“dependent on concrete and/or immediate experience”). These are the qualities that have been identified as indicators of the evidential use of \textit{estar}. Different modality encompasses variable contexts that are of great interest from both an empirical and theoretical perspective. On one hand, they are the locus of variation and change. On the other hand, they appear to defy attempts to present a unified formal analysis of copula distribution with adjectival predicates in Spanish. To reduce these contexts to a single variable is to miss a valuable opportunity to gain more information about the structure at hand. Consequently, I view semantic transparency less as a quantifiable variable than as a hypothesis about variation and change.

Perhaps the only true discourse-level variable that appeared in these studies was experience with the referent. Four studies included it in their analysis, and all found it to be a significant predictor in the selection of \textit{estar}. Experience with the referent was the only variable across all of the previous studies that explicitly targeted the evidential use of \textit{estar}. Geeslin and Guijarro-Fuentes (2008) provide some definition of the variable and its three categories of immediate, indirect and ongoing experience. They liken the classification of immediate experience to Delbecque’s (1997) notion of comparison between a speaker’s perception and the speaker’s preconceived expectations of the referent, and explain that “the comparison is not necessarily between two points in time, but rather between one’s expectations and what one actually finds in a given situation” (Geeslin & Guijarro-Fuentes, 2008, p. 373). This definition is reminiscent of mirativity, a category of meaning that is associated with evidentiality (see section 2.6.4 for discussion). In their contextualized preference task, Geeslin and Guijarro-Fuentes included a context that is associated with immediate experience with the referent (96).

(96) Raúl thinks the owner treated them badly because of their age. Paula does not agree and explains to him that she doesn’t look as young because she is a few years older than other students.

Paula: It can’t be because of that, I’m going to be 23 tomorrow.

Raúl: \textit{Ah, ¡qué vieja estás/eres!}
Oh, how old you are

(2008, p. 377)

Geeslin and Guijarro-Fuentes’ findings showed that experience with the referent was a significant predictor of estar for monolingual and Catalan and Galician bilingual Peninsular Spanish speakers, but not for Basque and Valencian bilinguals. Aguilar-Sánchez suggested that it was a significant predictor of estar in the Spanish of the multilingual community in Limón, Costa Rica, but did not quantify this variable. Díaz-Campos and Geeslin’s (2011) analysis of Caracas Spanish revealed that estar was strongly favoured in contexts of immediate experience. Indirect experience also favoured the selection of estar; ongoing experience slightly disfavoured the selection of estar. Díaz-Campos and Geeslin’s findings confirmed their predictions that the contexts that favoured estar involved both immediate experience, because it is associated with surprise or a reaction on the part of the speaker, and indirect experience, which “may favour estar because they include examples of subjective comments” (Díaz-Campos & Geeslin, 2011, p. 83). Brown and Cortés-Torres’s (2012) results indicated that immediate experience with the referent was significant in predicting the selection of estar in Puerto Rican Spanish. Estar was favoured in the immediate experience category and disfavoured in the indirect and ongoing categories. In sum, the category of immediate experience produced consistent predictions and results across studies, but the categories of ongoing and indirect experience showed variable results.

Given these divergent predictions and results, some clarification of the experience with the referent variable is needed. In this study, I claim that the evidential use of estar is distinct from other [estar + adjective] contexts. My solution is to adopt a definition and operationalization of the experience with the referent variable in terms of the categories outlined in formal approaches to evidentiality.

3.3.4 Summary of linguistic variables
Table 3.3 presents an overview of the linguistic variables examined in previous studies, and whether the variable was found to be significant in predicting the selection of *estar*.

Table 3.3. Summary of linguistic variables: Categories and findings

<table>
<thead>
<tr>
<th>Variable</th>
<th>Categories</th>
<th>Predictor of <em>estar</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Item-level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjective class</td>
<td>[Size, age, evaluative, physical appearance, etc.]</td>
<td>Significant</td>
</tr>
<tr>
<td>Resultant state adjective</td>
<td>[+/-Resultant state]</td>
<td>Mostly significant</td>
</tr>
<tr>
<td>Circumstantiality</td>
<td>[+/-Circumstantial]</td>
<td>Significant</td>
</tr>
<tr>
<td>Copulas allowed</td>
<td>[Ser only/Estar only/Both]</td>
<td>Significant</td>
</tr>
<tr>
<td><strong>Sentence-level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicate type</td>
<td>[Individual-level/Stage-level]</td>
<td>Mostly significant</td>
</tr>
<tr>
<td>Animacy of subject</td>
<td>[+/-Animate]</td>
<td>Not significant</td>
</tr>
<tr>
<td>Intensifier: <em>muy, más</em></td>
<td>[Present/Absent]</td>
<td>Significant</td>
</tr>
<tr>
<td>Temporal adverbs</td>
<td>[Present/Absent]</td>
<td>Significant</td>
</tr>
<tr>
<td>Norm/Frame of reference</td>
<td>[Individual/Class]</td>
<td>Mostly significant</td>
</tr>
<tr>
<td>Susceptibility to change</td>
<td>[+/-Susceptible to change]</td>
<td>Mixed findings</td>
</tr>
<tr>
<td><strong>Discourse-level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic transparency</td>
<td>[Different modality/Apparent synonymy/Choice not allowed/Clear difference]</td>
<td>Significant</td>
</tr>
<tr>
<td>Experience with the referent</td>
<td>[Immediate/Indirect/Ongoing]</td>
<td>Mostly significant</td>
</tr>
</tbody>
</table>
3.4 Results of previous studies: Extralinguistic variables

The extralinguistic variables evaluated in these studies include age, gender, level of education and linguistic status. The term *age* refers to a variable that measures copula use by speakers of different generations in real or apparent time, and which may be referred to as *generation* in some studies. I use *linguistic status* to refer to a collection of factors such as languages spoken, proficiency level in the languages spoken, and level of contact with the languages spoken.

3.4.1 Extralinguistic variables: Age

The previous studies I have discussed found many examples of interaction between age and other extralinguistic variables. Overall, the results were mixed as regards the variable of age, but two general trends emerged across the data for the varieties represented. The first is one in which the extension of *estar* appears as a change in progress, and the second is a pattern of gradual change or stable variation between conservative and innovative uses of *estar*.

The evidence from the Mexican-American, Mexican and Caribbean dialects suggests that a change is in progress in these varieties. In Silva-Corvalán’s (1986) data on Mexican-American bilinguals in Los Angeles, the percentage of the innovative use of *estar* was higher among younger speakers (age 15-29) than older speakers (age 30-65). Innovation was highest among speakers who immigrated to the U.S. during early childhood or who were born in the U.S. Given these differences between age groups, she concludes that the extension of *estar* is ongoing in this variety. Gutiérrez (1992) found that younger speakers (age 13-29) showed a higher frequency of the innovative use of *estar* than the two groups of older speakers (age 30-50 and 50+), who showed a lower frequency of the innovative use of *estar*. This suggests a change in progress in the monolingual Spanish of Morelia. Cortés-Torres (2004) discovered that age was only significant when analyzed in combination with the variable level of education in the monolingual Spanish of Cuernavaca. Speakers with less formal education showed the highest frequency of innovation with *estar*. The higher frequency of *estar* found in the middle (age 31-49) and
younger (age 18-30) age groups suggest a change in progress in this variety. Alfaraz’s (2012) investigation of the extension of *estar* in Cuban Spanish in both real and apparent time produced evidence in favour of a change in progress. Finally, Brown and Cortés-Torres’s (2012) findings indicated that age is a significant conditioning factor in the selection of *estar* in Puerto Rican Spanish, with younger speakers (age 20-29) showing the strongest tendency to favour *estar*. Although the data pattern appears to describe a change in progress, the authors point out that it may also reflect age-grading.

For the Caracas variety, the findings for the variable of age in the extension of *estar* indicated patterns of stable variation or gradual change. De Jonge (1993) found comparable percentages of the use of *estar* with age adjectives in the corpora of educated speakers in Caracas in the early 1970’s and again in the late 1980’s. In another study of Caracas Spanish, Díaz-Campos and Geeslin’s (2011) results revealed an interaction between age and socioeconomic class. Speakers from a lower socioeconomic background in both age groups (age 14-45 and age 46+) favoured *estar*, whereas the middle socioeconomic group demonstrated more conservative tendencies in both age groups. They encountered a difference between age groups for speakers from the upper socioeconomic class, where younger speakers favoured *estar* but older speakers disfavoured *estar*. Due to the uneven pattern of results across socioeconomic classes of younger speakers, the authors predict that the extension of *estar* will continue to advance gradually rather than quickly in this variety.

### 3.4.2 Extralinguistic variables: Gender

In hypothesizing about the role of gender in variation and change, many studies cite Labov’s (2001) generalization that women are more likely to lead change by adopting innovative forms, as long as the forms are not stigmatized. However, in the absence of any evidence regarding social awareness of the extension of *estar*, it is difficult to determine whether it is a stigmatized form, or make predictions on this basis about which gender will show a higher frequency of the

---

7 Age-grading is the phenomenon of change at the level of an individual speaker during his or her lifetime, rather than change at the level of the community (see for example Eckert, 1997; Wagner, 2012).
selection of *estar*. The findings of previous studies do not do much to clarify this picture either, as the majority of those that evaluated gender as a factor in *estar* usage discovered that it was not significant (Alfaraz, 2012; Brown & Cortés-Torres, 2012; Cortés-Torres, 2004; Díaz-Campos & Geeslin, 2011; Ortiz-López, 2000; Salazar, 2007; Silva-Corvalán, 1986).

Three studies that examined the variable of gender in the selection of *estar* obtained disparate results. Gutiérrez (1992) and Aguilar-Sánchez (2012) found greater rates of the use of *estar* among female speakers in Morelia, Mexico, and Limon, Costa Rica, respectively. Aguilar-Sánchez’s results revealed an interaction between gender and level of education in Limon, Costa Rica: women’s use of *estar* declined as their level of education rose, whereas men showed higher rates of the use of *estar* as their level of education increased. Geeslin and Guijarro-Fuentes (2007) obtained variable results for gender as a factor in *estar* selection across four groups of bilingual speakers and comparison groups of monolingual speakers of Iberian Spanish. In sum, with respect to variation and change with *estar*, most of the evidence from previous studies suggests that gender is not a significant predictor of the use of *estar*.

### 3.4.3 Extralinguistic variables: Level of education and linguistic status

The variables level of education and linguistic status – an umbrella term I use to refer to a collection of factors including languages spoken, proficiency level, age of acquisition, language dominance, and so on – are treated separately in some studies and as interacting in others. A common pattern emerges in the studies that investigate level of education without linking it to linguistic status. In both of the monolingual Mexican varieties of Morelia and Cuernavaca, speakers with less formal education showed higher rates of the innovative use of *estar* than speakers with more formal education (Cortés-Torres, 2004; Gutiérrez, 1992). Salazar (2007) also found that less formal education was a significant predictor of *estar* selection in her study of New Mexico and Colorado Spanish-English bilinguals.

The correlation between less formal education and the extension of *estar* becomes clearer when the interaction between level of education and linguistic status is taken into account. Silva-
Corvalán (1986) found that level of education alone was not significant in her analysis of Spanish-English bilinguals in Los Angeles, but rather, its interaction with proficiency level in Spanish was relevant to the extension of *estar* in this variety. A speaker’s lower proficiency level in Spanish was a significant predictor of more frequent use of innovative *estar* in his or her speech. In particular, early childhood bilinguals displayed a much greater tendency towards the extension of *estar* than speakers who acquired English during adolescence or adulthood. Silva-Corvalán suggests two possible interpretations for this finding. One possibility is that early contact between two languages leads to the preference for a single form in the language whose system contains two forms. The second and more likely possibility, in her opinion, is that early contact between two languages has the effect of accelerating change in the subordinate language. She also speculates that reduced exposure to full-fledged varieties of the language, including access to formal institutions where conservative forms are upheld, works to magnify the effect of language contact and brings about even more rapid diffusion of the change in the community.

Ortiz-López (2000) expands upon Silva-Corvalán’s findings in his investigation of the extension of *estar* among Spanish-English bilinguals and monolinguals in Puerto Rico. At first glance, his results reveal a surprising contrast to those of Silva-Corvalán’s study: monolingual speakers showed higher rates of the innovative use of *estar* than balanced bilinguals. This finding appears to contradict Silva-Corvalán’s hypothesis about the role of language contact in accelerating the change. However, Ortiz-López hypothesizes that it is the interaction between level of formal education in Spanish and proficiency level in English that is responsible for the outcomes he encounters. Specifically, a higher level of formal education in Spanish is a characteristic of the bilingual speakers in his study. This factor has a neutralizing effect on the potential for change that contact with English represents.

The findings of Aguilar-Sánchez’s (2012) study of copula use among monolingual, bilingual and multilingual speakers in Limón, Costa Rica, matched many of the results of previous studies with respect to the variables of level of education and linguistic status. The author identified two varieties of Spanish in the region: a monolingual variety and a bilingual/multilingual variety. Speakers of the monolingual variety showed a decrease in their use of *estar* as their level of formal education in Spanish increased. Bilingual/multilingual speakers displayed higher rates of *estar* selection as their level of formal education in English rose, and greater use of *estar* overall.
as compared with monolingual speakers. These results bolster the hypothesis that the extension of *estar* is associated with reduced access to formal education in Spanish (Ortiz-López, 2000; Silva-Corvalán, 1986).

Geeslin and Guijarro-Fuentes (2008) posed the question of whether the increased pace of the change affecting *estar* could be attributed to language contact in general or contact with English in particular. The authors addressed the role of factors such as the cognitive demands of bilingualism, typological similarity of the contact language and language dominance in the change affecting the copula system in Spanish. The results for the frequency of *estar* selection in Spain among four bilingual groups (speakers of Catalan, Valencian, Galician and the typologically distinct Basque language) were significantly different from the results for the monolingual group, in both directions. The rates of *estar* selection among Catalan-Spanish and Valencian-Spanish bilingual groups were lower than those of the monolingual Spanish group, whereas the Basque-Spanish and Galician-Spanish bilinguals showed higher rates of the selection of *estar* than monolinguals. The authors interpret these findings to mean that the effects of language contact on the copula system are not always linked to the extension of *estar*. Geeslin & Guijarro-Fuentes also hypothesize that language dominance may play a role in the increased use of *estar* based on their finding that Galician-Spanish bilinguals displayed the highest rates of *estar* selection of all groups, and this group also reported feeling less comfortable in Spanish than other groups. However, they dismiss the hypothesis that the cognitive demands of bilingualism play a role in the increased use of *estar*, pointing to variation in copula selection even among monolinguals as counterevidence. The finding that the Basque-Spanish bilinguals showed lower rates of *estar* usage than Galician-Spanish bilinguals suggests that a typological difference between languages, as in the situation of Spanish-English bilinguals in previous studies, does not necessarily speed up the pace of change. Aguilar-Sánchez’s findings also support Geeslin & Guijarro-Fuentes’s claim that it is not language contact in general but contact with English in particular that leads to an acceleration of the change affecting *ser* and *estar*. Why English, and not other languages, should be linked to an acceleration of the process of change with *ser* and *estar* in Spanish is an interesting question, but it is one which is not addressed in these studies.
3.4.4 Summary of extralinguistic variables

Several general trends are expressed by the results for each of the extralinguistic variables. The data on the variable of age point to a change in progress in some varieties and a situation of gradual change or stable variation in others. Gender, for the most part, was not significant as a predictor of *estar* in the studies reviewed here. By contrast, level of formal education was mostly found to be a significant predictor of *estar* across these studies. Linguistic status (including languages spoken, proficiency level, age of acquisition and language dominance) was also suggest a significant predictor of the use of *estar*. Table 3.4 provides an overview of the extralinguistic variables analyzed in each of the previous studies, and whether the variable was found to be significant in predicting the selection of *estar*.

Table 3.4. Summary of extralinguistic factors: Variables and findings

<table>
<thead>
<tr>
<th>Study</th>
<th>Age</th>
<th>Gender</th>
<th>Education</th>
<th>Linguistic Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silva-Corvalán 1986</td>
<td>√*</td>
<td>√</td>
<td>√</td>
<td>√*</td>
</tr>
<tr>
<td>Gutiérrez 1992</td>
<td>√*</td>
<td>√</td>
<td>√*</td>
<td>X</td>
</tr>
<tr>
<td>de Jonge 1993</td>
<td>√</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ortiz-López 2000</td>
<td>X</td>
<td>X</td>
<td>√*</td>
<td>√*</td>
</tr>
<tr>
<td>Cortés-Torres 2004</td>
<td>√</td>
<td>√</td>
<td>√*</td>
<td>X</td>
</tr>
<tr>
<td>Salazar 2007</td>
<td>X</td>
<td>√</td>
<td>√*</td>
<td>X</td>
</tr>
<tr>
<td>Geeslin &amp; Guijarro-Fuentes 2007</td>
<td>X</td>
<td>√*(some groups)</td>
<td>X</td>
<td>√*</td>
</tr>
<tr>
<td>Geeslin &amp; Guijarro-Fuentes 2008</td>
<td>√*(1 group)</td>
<td>√*(some groups)</td>
<td>√*</td>
<td>√*</td>
</tr>
<tr>
<td>Diaz-Campos &amp; Geeslin 2011</td>
<td>√*</td>
<td>√</td>
<td>√*</td>
<td>X</td>
</tr>
<tr>
<td>Aguilar-Sánchez 2012</td>
<td>√*</td>
<td>√*</td>
<td>√*</td>
<td>√*</td>
</tr>
<tr>
<td>Alfaraz 2012</td>
<td>√*</td>
<td>√</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Brown &amp; Cortés-Torres 2012</td>
<td>√*</td>
<td>√</td>
<td>X</td>
<td>√</td>
</tr>
</tbody>
</table>

√=variable included  X=variable not included  *=significant predictor of *estar*
The extension of *estar* into contexts previously occupied by *ser* is a change that has long been documented as part of the evolution of the copula system in Spanish, and it is one that continues in the present day in some varieties. Several recent empirical studies of variation and change in copula selection in Spanish have investigated diverse populations and varieties. Methods of analysis have included the binary classification of innovative versus prescriptive uses of *estar*, multivariate analysis, and a usage-based approach. The results indicate that while the change is subject to strong conditioning of linguistic factors, extralinguistic factors such as age, level of education and linguistic status also play a role.

Of the linguistic variables tested in previous studies, the strongest predictors of the use of *estar* were adjective class, the presence of temporal adverbs in the context, norm/frame of reference and experience with the referent. The adjective classes that occurred consistently in combination with *estar* across the varieties studied included age, size, physical appearance and evaluation. Evaluative adjectives (e.g. *bueno* ‘good’, *bonito* ‘pretty’) appeared in some contexts that were classified as innovative uses of *estar* in previous studies but which I identify as evidential uses of *estar*. A usage-based analysis (Brown & Cortés-Torres, 2012) provided a qualitative overview of the lexical diffusion of the innovative use of *estar*. Many of the adjectives that acted as central members of the semantic clusters of related adjectives that appeared in the pragmatic [*estar* + adjective] construction were evaluative adjectives, suggesting that at least some of these innovative contexts were evidential uses.

An overview of the role of the variable of age in *ser/estar* selection presents several possible trends across varieties: ongoing change, gradual change and stable variation. In broad terms, change that involves the semantic bleaching of the *ser/estar* contrast is in progress in Caribbean, Mexican and U.S. dialects (Alfaraz, 2012; Brown & Cortés-Torres, 2012; Cortés-Torres, 2004; Gutiérrez, 1992; Ortiz-López, 2000; Salazar, 2007; Silva-Corvalán, 1986). Silva-Corvalán characterizes the path of change in Spanish-English bilinguals in Los Angeles as a loss of selectional restrictions on *estar*. When variation appears in the speech of individuals, this reflects
an alternation between two options (or two stages) in the grammatical system of the speaker, in which one option is more restricted and the other option is less restricted in its selection with *estar*. Studies of the Caracas variety suggest that the use of *ser* and *estar* with adjectives there is more aptly described either in terms of gradual change or stable variation between conservative and innovative uses of *estar* (de Jonge, 1993; Díaz-Campos & Geeslin, 2011). De Jonge’s analysis contrasting the Caracas and Mexico City dialects found that both showed internal stability but contrasted in terms of the stage of the change affecting *estar* that they represented.

Other extralinguistic variables evaluated in previous studies included gender, level of education and linguistic status. For the most part, gender was not found to be a significant predictor of the use of *estar* in these previous studies. However, there was evidence of interaction between level of education and factors relating to linguistic status, such as bilingualism, proficiency level, age of acquisition, contact language and so on. The greatest tendency towards the extension of *estar* appeared in bilingual varieties in contact with English. Reduced exposure to formal education in Spanish in bilingual varieties was also associated with the increased use of *estar*.

The findings of these previous studies indicate that there is a need to differentiate between cases of semantic bleaching between the copulas and the evidential use of *estar*, where *estar* is selected as a strategy for the expression of the speaker’s perspective. Geeslin and Guijarro-Fuentes (2008) note that a quantitative analysis of the experience with the referent variable requires a robust data set, which suggests that the use of *estar* to express or emphasize the speaker’s viewpoint is an infrequent use of this verb. Nonetheless, I claim that the evidential use of *estar* expresses a distinct meaning, and this meaning cannot be associated with any of the other variables that have been proposed to account for *ser/estar* selection. In the chapter that follows, I discuss my hypothesis in further detail and present the methodology I use to test my predictions about the evidential use of *estar*. 
Chapter 4
Hypothesis and methodology

4.1. Hypothesis and predictions

4.1.1. Hypothesis

Building on previous analyses of the evidential use of *estar*, (Camacho, 2012a, 2012b; Escandell-Vidal & Leonetti, 2002; Roby, 2009), I propose to test the hypothesis that the [*estar + adjective*] context may express evidential meaning. The evidential use of *estar* co-exists with, but can be distinguished from, other interpretations of *estar* with adjectival predicates, such as temporally-limited, change-of-state, and resultative meanings. I view this use of *estar* as an evidentiality strategy that expresses a language universal: the encoding of speaker perspective, specifically the encoding of source of information. The evidential use of *estar* is optional and relies on the contrast with *ser* in the dual copula system in Spanish. *Estar* expresses evidentially marked discourse (2); *ser* expresses neutral discourse (3).

(2)  

a.  *Este vino está riquísimo.*

‘This wine is\textsubscript{ESTAR} delicious\textsubscript{INT}.’

b.  *¿Viste la nueva película de Almodóvar? Está buena.*

‘Did you see the new Almodovar film? It is\textsubscript{ESTAR} good.’

c.  *Están muy lindos los cachorros del vecino.*

‘The neighbor’s puppies are\textsubscript{ESTAR} cute.’

---

8 Semantic change is typically characterized by the co-existence of older and newer meanings of a form, a process known as “layering” (Hopper, 1991, p. 22).
(3)  a.  *Este vino es riquísimo.*

   ‘This wine is delicious.’

   b.  ¿*Viste la nueva película de Almodóvar? Es buena.*

   ‘Did you see the new Almodovar film? It is good.’

   c.  *Son muy lindos los cachorros del vecino.*

   ‘The neighbor’s puppies are cute.’

What sets the sentences in (2) apart from those in (3) is that *estar* encodes the speaker’s experience as the source of the information expressed in the sentence. The sentences with *ser* in (3) are neutral as to their information source: from the speaker’s perspective, the information source is irrelevant.

This meaning is illustrated more clearly by the contradiction that arises if the speaker’s knowledge of the information source is denied. I consulted three (non-linguist) native speakers of Buenos Aires Spanish for the contradiction tests in (97-98), and the consultants reached 100% agreement on these judgments. If an additional clause is added in which the speaker’s experience is denied, the *estar* predication is contradicted (97), whereas the corresponding *ser* predication (98) is not.

(97)  a.  *El vino está rico (pero yo no lo probé).*

   ‘The wine is delicious (but I didn’t taste it).’

   b.  *La película está buena (pero yo no la vi).*

   ‘The film is good (but I didn’t see it).’

   c.  *Los cachorros están lindos (pero yo no los vi).*

   ‘The puppies are cute (but I didn’t see them).’

(98)  a.  *El vino es rico (pero yo no lo probé).*
‘The wine is SER delicious (but I didn’t taste it).’

b. La película es buena (pero yo no la vi).

‘The film is SER good (but I didn’t see it).’

c. Los cachorros son lindos (pero yo no los vi).

‘The puppies are SER cute (but I didn’t see them).’

The variation and change in the distribution and interpretation of *ser* and *estar* is characterized by ongoing reanalysis, and the evidential use of *estar* is part of the evolution of the copula distinction. Crucially, it is distinct from the semantic bleaching that has been documented for some varieties of Spanish in contact with English, in which *ser* and *estar* have become interchangeable in combination with adjectival predicates. Previous studies of variation and change have not distinguished evidential uses from cases of semantic bleaching of *estar*, grouping them together under the heading of “innovation”. One of the objectives of the current study is to show that evidential readings of *estar* predications, although infrequent, express a distinct meaning. As such, they represent a different outcome of the process of change from the loss of meaning characterized by semantic bleaching.

My second objective is to examine the conditioning factors of the contexts of the evidential use of *estar*. The analysis includes both internal and external factors. My predictions in regards to internal or linguistic factors are informed by theoretical proposals and typological generalizations from the literature on evidentiality. Generally, I investigate whether speakers show a difference in patterns of selection of *ser* and *estar* depending on whether evidence is present in the context. A more specific question relates to the three main categories of evidence type: direct, inferred, and reported evidence, and which of these are expressed by [*estar* + adjective]. Another line of inquiry concerns sensory evidence in particular, and asks whether speakers show a difference in preferences of *ser* and *estar* according to the sense that is involved in the perception of evidence (visual/auditory/olfactory/tactile/gustatory).

I also predict that adjective type plays a role in the availability of evidential readings. I expect to find that evidential meaning appears in the presence of evaluative adjectives. Following previous
studies of the semantic extension of *estar* (see chapter 3 for discussion), I understand the
category of evaluative adjectives to be based on the lexical semantic classification of adjectives
identified by Dixon (1977, 2004). In my definition, an evaluative adjective is a qualificational
adjective whose primary lexical content is a positive or negative evaluation of a noun or
proposition (prototypical examples are *bueno* and *malo*, ‘good’ and ‘bad’). Finally, I predict that
some of the variables of previous studies can be reframed in terms of a theory of gradability. I
expect that some of the variation in the use of *ser* and *estar* with adjectives can be accounted for
in terms of an analysis of the quantization of stative predicates.

With respect to social variables, I consider the roles that age and level of formal education play
in the use of *estar* as an evidentiality strategy. I examine whether speakers of different age
groups (younger vs. older) and educational backgrounds (less vs. more formal education) pattern
differently in their selection of *ser* and *estar* with adjectives based on the presence of evidence in
the context. Previous studies have found that younger speakers and less educated speakers lead
the change in the expansion of the innovative use of *estar* in Mexican, Caribbean and U.S.
varieties of Spanish (Alfaraz, 2012; Brown & Cortés-Torres, 2012; Cortés-Torres, 2004;
Gutiérrez, 1992; Ortiz-López, 2000; Salazar, 2007; Silva-Corvalán, 1986). However, studies of
Caracas Spanish have identified a situation of stable variation/gradual change in the use of *ser*
and *estar* with adjectives in that dialect (de Jonge, 1993; Díaz-Campos & Geeslin, 2011). In this
study, I explore the roles of the extralinguistic variables of age and level of education in the
evidential use of *estar* in order to characterize the social dimension of the variation and/or
change in the speech community of Buenos Aires.

### 4.1.2 The evidential use of *estar* in Buenos Aires Spanish

To the best of my knowledge, no previous studies of variation and change in the use of *ser* and
*estar* with adjectives have looked at the Buenos Aires variety. I propose to investigate the
selection of *ser* and *estar* with adjectives in this dialect in order to study the evidential use of
*estar* and identify the conditioning factors present in its contexts of use. The evidential use of
*estar* is an optional and infrequent use of this verb, but one which should be available in all dialects. The fact that this use appears as far back as the 16th century in literary texts is an argument in favour of the existence of this meaning. A finding of the evidential use of *estar* in the Buenos Aires variety strengthens the case.

Buenos Aires Spanish is ideal for my purposes, given that it is not affected by the kinds of language contact factors that have the potential to obscure the main questions of this study. As we have seen, documentation of varieties of Spanish in contact with English (i.e., U.S., Mexico, and Caribbean dialects) reveals that the semantic extension of *estar* is more advanced in these varieties. The bleaching of the semantic contrast between *ser* and *estar* may be accelerated by contact with English. Buenos Aires Spanish has had minimal contact with English compared with the varieties undergoing change, which suggests that the semantic bleaching observed in other varieties may be less advanced in this variety. The evidential use of *estar* can sometimes be ambiguous with other interpretations of *estar* clauses, including innovative uses of *estar* (see section 3.3.1 for discussion and examples). Although it is not realistic to expect to completely eliminate ambiguous contexts, ambiguity can be minimized by choosing a variety where it is not robust.

Other varieties of Spanish have acquired evidential extensions to the tense system, in cases where the contact language expresses evidentiality through grammatical marking. For instance, some varieties of Andean Spanish in contact with Quechua and Aymara have developed evidential meaning extensions on the perfect tenses (de Granda, 1994; Escobar, 1997; Klee & Ocampo, 1995; Sánchez, 2004). Also, the expression of evidential contrasts via the past tenses has been documented in Paraguayan Spanish in contact with Guarani (Palacios, 2008). The consequences for the evidential use of *estar* in these varieties are unknown, but if another form is available for the expression of evidential meaning, the evidential use of *estar* may be less robust. This is another potential problem that I hope to avoid through the choice of Buenos Aires Spanish as the variety in which to investigate *estar’s* evidential meaning.

For all of these reasons, the Buenos Aires dialect is an optimal variety for isolating the evidential use of *estar*, for investigating the roles of conditioning factors on this use, and for adding data on
an as-yet undocumented variety to our understanding of variation and change in the use of *ser* and *estar* with adjectives across the Spanish-speaking world.

### 4.2 Methodology

To test the hypotheses outlined in the previous section, I conducted a three-part experiment, consisting of a written contextualized preference task, a controlled oral elicitation task, and a guided interview. The oral tasks (interview and elicitation) were completed prior to the contextualized preference task to prevent any potential priming effects of the presence of the target structure in the written task. All participants completed all three tasks, and all components of the experiment were conducted in Spanish.

#### 4.2.1 Contextualized Preference Task

#### 4.2.1.1 Background

The contextualized preference task (CPT) was a written questionnaire, and it is a type of task that was used previously to test the use of *ser* and *estar* with adjectival predicates in Peninsular Spanish by Geeslin and Guijarro-Fuentes (2008). The main objective of this task was to investigate which, if any, of the three broad semantic parameters of evidentiality (direct/inferred/reported) are encoded by *estar* when it is used as an evidentiality strategy (see section 2.6.3 for discussion of this tripartite division of evidence types). This task tested whether speakers used *estar* as an evidentiality strategy for none, some, or all, of these categories. The task also examined whether the social variables of age and socioeconomic status played a role in participants’ use of *estar* as an evidentiality strategy.
4.2.1.2 Method

The CPT was a written task, and differed from the oral tasks in two important ways. One of the undesired effects of written tasks is that they can induce a more formal register or a more prescriptive response from participants, given that written communication often requires a more formal register and the format of the task may be reminiscent of an academic test. This is the main disadvantage when using a written task. However, there are also advantages. In an interview, or even in a more focused elicitation task, the researcher can only prompt participants towards the use of the target structures. In production tasks, there is no guarantee that the participants will produce the target contexts. Participants may not use them at all, or may use them with such low frequency that it is impossible to produce statistically significant results with the number of tokens obtained. By contrast, a written task such as the CPT gives the researcher full control over the language that is used in the task. The researcher has the freedom to create test conditions that include specific structures or lexical items, as well as to make up a robust enough set of test conditions to produce results that will have statistical significance. In short, the written task is less natural, but more targeted, than the oral tasks. This combination of written and oral tasks has been chosen with the goal of providing a fuller picture of the structure in question than any one type of task would on its own.

The CPT contained 35 items and presented a story featuring a conversation between two characters, Adriana and Juan Pablo, who go out for dinner in Buenos Aires and discuss everyday topics such as food, shopping, travel, friends and family, parties, and so on. Each item presented an introductory two- to three-sentence context, followed by a minimal pair of sentences containing either ser or estar in combination with an adjective. Participants selected which of the two sentences they preferred. A second component to each item appeared below the minimal pair of sentences. It gave participants a second chance to make a choice between the sentences, and it read: “If both sentences are acceptable, and you were going to say this, how would you say it? ___ The first ___ The second ___ I can’t decide.” This component was included for the purpose of eliminating as many ambiguous answers as possible. A complete breakdown of items in the CPT appears in Table 4.1.
Table 4.1. Design: Contextualized Preference Task

<table>
<thead>
<tr>
<th>Type of item</th>
<th># of items</th>
<th>Total items</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: evidential: direct</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>B: neutral: direct</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>C: evidential: inferred</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>D: neutral: inferred</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>E: evidential: reported</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>F: neutral: reported</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>G: distractors</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>H: practice items</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td><strong>Total items</strong></td>
<td></td>
<td><strong>35</strong></td>
</tr>
</tbody>
</table>

The four items in Condition A tested the use of *estar* to encode direct (including visual and non-visual sensory) evidence. The same four adjectives were tested in the corresponding evidentially neutral contexts of Condition B. Condition C presented four items that tested the use of *estar* for the inferred (including inference and assumption) evidence parameter, and Condition D again contrasted the same adjectives in neutral contexts. Finally, Condition E was designed using four adjectives in contexts of reported (including hearsay and quotative) evidence, and these items were matched by four neutral items featuring the same adjectives in Condition F. Table 4.2 presents the adjectives and subjects that appeared in each condition.
Table 4.2. Adjectives by Condition: Contextualized Preference Task

<table>
<thead>
<tr>
<th>Evidence type</th>
<th>Adjectives</th>
<th>Subject nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition A:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct evidence</td>
<td><em>grande, lindo, seco, amable</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(big, cute, dry, friendly)</td>
<td><em>patio, valija, lomo, mozo</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(patio, suitcase, steak, waiter)</td>
</tr>
<tr>
<td>Condition B:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td><em>grande, lindo, seco, amable</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(big, cute, dry, friendly)</td>
<td><em>casa, globo, Madrid, tío</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(house, balloon, Madrid, uncle)</td>
</tr>
<tr>
<td>Condition C:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferred evidence</td>
<td><em>feo, chico, rico, tímido</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ugly, small, delicious, shy)</td>
<td><em>barrio, cámara, comida, tía</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(neighbourhood, camera, food, aunt)</td>
</tr>
<tr>
<td>Condition D:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td><em>feo, chico, rico, tímido</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ugly, small, delicious, shy)</td>
<td><em>letra, mandolina, queso, Sergio</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(lyrics, mandolin, cheese, Sergio)</td>
</tr>
<tr>
<td>Condition E:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reported evidence</td>
<td><em>espantoso, barato, bueno, alto</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(disgusting, cheap, good, tall)</td>
<td><em>restaurante, zapatos, película, novia de Antonio</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(restaurant, shoes, film, Antonio’s girlfriend)</td>
</tr>
<tr>
<td>Condition F:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td><em>espantoso, barato, bueno, alto</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(disgusting, cheap, good, tall)</td>
<td><em>vino, cámara, escuela, tía</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(wine, camera, school, aunt)</td>
</tr>
</tbody>
</table>

The following examples show a test condition (99) and a neutral condition (100) using the same adjective to test for direct evidence.

(99) **Condition A: Direct evidence [Evidential]**

*Deciden que quieren sentarse afuera. Se encuentran en una calle donde algunos restaurantes tienen patios. Adriana se para al lado de la puerta de un restaurante que se llama Las Flores.*

*Adriana: ¿Qué te parece este lugar?*

*Juan Pablo:*
A. Bien, el patio es grande.

B. Bien, el patio está grande.

‘They decide that they want to sit outside. They find themselves in a street where some of the restaurants have patios. Adriana stops beside the door of a restaurant called Las Flores.

Adriana: What do you think of this place?

Juan Pablo:

A: Well, the patio es big.

B: Well, the patio está big.’

(100) Condition B: [Neutral]

Adriana le dice a Juan Pablo que tiene una noticia emocionante. Le cuenta que acaba de comprar el pasaje a Barcelona para el mes próximo.

Juan Pablo: ¡Buenísimo! ¿Dónde te vas a quedar allí?

Adriana:

A. Me voy a quedar en lo de mi tía; su casa está grande.

B. Me voy a quedar en lo de mi tía; su casa es grande.

‘Adriana tells Juan Pablo that she has exciting news. She tells him that she has just bought her ticket to Barcelona for the following month.

Juan Pablo: Great! Where are you going to stay there?

Adriana:

A: I’m going to stay at my aunt’s place; her house está big.

B: I’m going to stay at my aunt’s place: her house es big.’
These examples are of a test condition (101) and a neutral condition (102) using the same adjective to test for inferred evidence.

(101) Condition C: Inferred evidence [Evidential]

Adriana le pregunta a Juan Pablo si sabe la marca y el modelo de cámara que quiere. Él le dice que estaba leyendo en internet sobre una cámara nueva que tiene todas las características que él quiere.

Adriana: ¡Espero que quepa en la valija!

Juan Pablo:

A. Creo que la cámara es chica porque cabe en el bolsillo.

B. Creo que la cámara está chica porque cabe en el bolsillo.

‘Adriana asks Juan Pablo if he knows the brand and the model of the camera that he wants. He tells her that he was reading on the Internet about a new camera that has all the qualities he wants.

Adriana: I hope it fits in the suitcase!

Juan Pablo:

A. I think the camera is small because it fits in your pocket.

B. I think the camera is small because it fits in your pocket.’

(102) Condition D: [Neutral]

Desde donde están sentados en el patio, pueden oír la música en vivo de algún lugar cercano. Juan Pablo le pregunta a Adriana si sabe el nombre del instrumento que se toca, y ella le dice que suena como una mandolina.

Juan Pablo: ¿Qué es una mandolina?
Adriana:

A. Se parece a una guitarra, pero está chica.

B. Se parece a una guitarra, pero es chica.

‘From where they are sitting on the patio, they can hear live music from somewhere nearby. Juan Pablo asks Adriana if she knows the name of the instrument that is playing, and she tells him it sounds like a mandolin.

Juan Pablo: What is a mandolin?

Adriana:

A. It’s like a guitar, but it is ESTAR small.

B. It’s like a guitar, but it is SER small.’

Here are examples of a test condition (103) and a neutral condition (104) using the same adjective to test for reported evidence.

(103) Condition E: Reported evidence [Evidential]

Adriana le pregunta a Juan Pablo quién está invitado a la fiesta. Juan Pablo le dice que Antonio va a traer a su novia nueva.

Adriana: ¿Sabés algo de ella?

Juan Pablo:

A: Dos cosas: Antonio dijo que se conocieron bailando, y que está alta.

B: Dos cosas: Antonio dijo que se conocieron bailando, y que es alta.

‘Adriana asks Juan Pablo who is invited to the party. Juan Pablo tells her that Antonio is going to bring his new girlfriend.

Adriana: Do you know anything about her?
Juan Pablo:

A: Two things: Antonio said that they met dancing, and that she is ESTAR tall.

B: Two things: Antonio said that they met dancing, and that she is SER tall.’

(104) Condition F: [Neutral]

Juan Pablo dice que no sabía que Adriana tenía una tía en Barcelona. Adriana le dice que se llama Beti, es la hermana menor de su madre, y que todo el mundo dice que Beti y Adriana se parecen.

Juan Pablo: ¿Cómo se parecen?

Adriana:

A: Ella es alta, como yo.

B: Ella está alta, como yo.

‘Juan Pablo says that he didn’t know that Adriana had an aunt in Barcelona. Adriana tells him that her aunt is named Beti, is her mother’s younger sister, and that everyone says that Beti and Adriana look alike.

Juan Pablo: How do you look alike?

Adriana:

A: She is SER tall, like me.

B: She is ESTAR tall, like me.’

The first two items in the test instrument were practice items, and the order of the remaining 33 items was pseudo-randomized. The structures presented in the nine distractor and two practice items were chosen for their previously documented vulnerability to variation and/or change across dialects or in the Buenos Aires dialect in particular (see Lipski, 1994). Table 4.3 presents the structures and lexical items that appeared in the distractor and practice items.
Table 4.3. Distractors and Practice Items: Contextualized Preference Task

<table>
<thead>
<tr>
<th>Structure</th>
<th>Lexical items</th>
<th>Number of items</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>de-que-ismo:</em> insertion/omission of <em>de</em></td>
<td><em>acordarse (de) que,</em></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td><em>preocuparse (de) que</em></td>
<td></td>
</tr>
<tr>
<td>preterite vs. imperfect tense</td>
<td><em>salir, durar</em></td>
<td>2</td>
</tr>
<tr>
<td>pluralization of <em>haber</em></td>
<td><em>va/van a haber, había/habian</em></td>
<td>2</td>
</tr>
<tr>
<td>subjunctive vs. indicative mood</td>
<td><em>es importante que,</em> no creo que*</td>
<td>2</td>
</tr>
<tr>
<td>preterite vs. present perfect tense</td>
<td><em>cambiar, ver</em></td>
<td>2</td>
</tr>
<tr>
<td>impersonal <em>se</em> vs. passive <em>se</em></td>
<td><em>se vende/se venden</em></td>
<td>1</td>
</tr>
</tbody>
</table>

There were two versions of the test instrument. The order of the items was the same but the two versions were counterbalanced with respect to the order in which *ser* and *estar* were presented in the minimal pairs. Within each version, the order of presentation of *ser* and *estar* alternated from each item to the next (for test condition items). The complete set of stimuli appears in Appendix A.

4.2.1.3 Participants

Forty-five native speakers of Buenos Aires Spanish were recruited to participate in the study, which took place over a six-week period during March and April 2012. The researcher recruited participants by email, on Facebook, and through personal contacts. All participants were either monolinguals or late bilinguals with no more than an intermediate level of proficiency in their other language(s). Information regarding language history, age and socioeconomic status was
obtained through participants’ voluntary disclosure on the language and personal background questionnaire, which appears in its entirety in Appendix B.

Previous studies have found that the social variables of age and socioeconomic status, but not gender, play a role in rates of the selection of ser and estar (see Cortés-Torres, 2004; Geeslin & Guijarro-Fuentes, 2008; Gutiérrez, 1992; Silva-Corvalán, 1986). Taking these findings as a point of departure, I selected participants according to age and socioeconomic status, but not according to gender. To test the role of the social variable of age, and to determine whether a change may have taken place over time, speakers from two age groups were recruited: younger speakers and older speakers. Younger speakers were defined as between the ages of 18 and 30, and older speakers were defined as age 50 and over. Speakers from two distinct socioeconomic backgrounds were recruited to investigate the role of socioeconomic status in rates of ser/estar selection. During recruitment, socioeconomic status was determined by a combination of factors: level of formal education, neighbourhood of residence, and occupation. However, over the course of data transcription and analysis, I determined that level of education was the most reliable factor in predicting socioeconomic status. Neighbourhood of residence was unreliable due to the diversity of income levels that coexist in close proximity in the urban environment of Buenos Aires, making it difficult to classify some neighbourhoods clearly in a single category. Occupation was also unreliable as a determining factor in socioeconomic status, as there were several recurring occupations that did not neatly fall into any category, such as retiree, student, musician, and homemaker. Thus, although information about participants’ neighbourhoods of residence and occupations was collected, the single factor that was ultimately used to determine socioeconomic status was the participant’s highest level of formal education obtained. There were two categories of level of education. One group was made up of participants who had obtained less formal education, which was defined as having completed high school or less (primary, secondary, or incomplete secondary education). Participants in the second group had obtained more formal education, which was defined as having completed one or more years of post-secondary education (vocational school or university). A brief synopsis of the relevant participant data appears in Table 4.4, and a complete inventory of participant data can be found in Appendix C.
Table 4.4. Overview of participant data by age and level of education

<table>
<thead>
<tr>
<th>Age range</th>
<th>Highest level of formal education completed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primary</td>
<td>Secondary</td>
</tr>
<tr>
<td>18-30</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>50-96</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Sub-total</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>24</td>
</tr>
</tbody>
</table>

4.2.1.4 Procedure

The CPT was administered on paper, and participants completed it after the interview and elicitation tasks. Presenting the tasks in this order eliminated the possibility of priming effects from the language in the written comprehension task on the oral production tasks. I briefly instructed participants on how to conduct the task and assisted participants with the first two practice items to ensure that participants understood the instructions correctly. Participants read the item and indicated which sentence they preferred, and if they chose both responses, I asked them to indicate which of the two sentences they would say themselves. I also invited them to make a note of any comments or additional responses by writing them directly on the paper questionnaire. Participants were permitted to reread the contexts as many times as they wished, and there was no time limit to the task. In two cases where an older participant had limited vision, a family member who was a native speaker of Buenos Aires Spanish was recruited to read aloud the entire questionnaire, and I wrote down the participant’s preferences and any additional responses or comments.
4.2.2 Controlled Elicitation Task

4.2.2.1 Background

The controlled elicitation task was a focused oral task that offered a precise approach to examining \textit{ser/estar} selection in evidential contexts. The aim of this task was to elicit production of the target structures (\textit{ser/estar} + adjective) in contexts where direct, sensory evidence was present, and to contrast this production with elicited production in contexts where direct, sensory evidence was not present. Variation across groups was later analyzed to consider the role of the social variables of age and socioeconomic status in participants’ choice of \textit{ser/estar} in evidential and neutral contexts.

4.2.2.2 Method

Participants were presented with sensory stimuli for each of the five senses in the following order: sight > hearing > smell > touch > taste. I presented the items to the participants and prompted them to express a spoken reaction to or impression of each stimulus. They were then asked to contrast the stimulus with a similar item in their previous experience. The expressions used in the prompts were as follows (105).

\begin{enumerate}
\item[(105)] a. Direct evidence stimuli:
\[
\text{¿Qué le/te parece ____________?}
\]
\text{‘What do you think of ____________?’}

b. Recalled experience stimuli:
\[
\text{¿Y qué me dice/decís de ________________?}
\]
\end{enumerate}
‘What can you tell me about _____________________?’

This contrast between direct evidence and recalled experience stimuli was designed to explore whether there is a difference in the selection of *ser* and *estar* with adjectival predicates in contexts where direct evidence is present (experimental stimuli) as compared with contexts where direct evidence is not present (recalled experience). Although elicited production tasks are widespread in linguistic research, I am not aware of any previous use of this design. Given the ambiguity of meaning that is present in many contexts of the evidential use of *estar*, one of the challenges in studying this structure is the researcher’s lack of access to the intentions of the speaker. This task was designed to reduce potential ambiguity of meaning by controlling the presence of evidence in the discourse context. It has the advantage of eliciting spontaneous production, like the interview task, but also offers the benefit of a tighter focus on the target structure and some control over relevant aspects of the discourse context, like the contextualized preference task.

The items chosen for experimental stimuli presented a variety of visual, auditory, olfactory, tactile, and gustatory input, in order to elicit as wide a range as possible of adjectival predicates in participants’ responses. Visual stimuli consisted of eight 20 x 25 cm photographic images (plus one additional image used for practice), each printed on individual sheets of A4-size paper and laminated. The images were obtained at no cost from the digital photo archive *FreeDigitalPhotos.net* during the months of January and February, 2012. The complete set of images and credits appear in Appendix D. Table 4.5 presents the set of contrasting direct evidence and recalled experience visual stimuli.
Table 4.5. Controlled Elicitation Task: Visual Stimuli: Images versus Recalled Images

<table>
<thead>
<tr>
<th>Stimulus code</th>
<th>Stimulus type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice</td>
<td>Direct</td>
<td>Boy dressed as a businessman</td>
</tr>
<tr>
<td>I1d</td>
<td>Direct</td>
<td>Landscape of clouds, sky and mountains</td>
</tr>
<tr>
<td>I1r</td>
<td>Recalled</td>
<td>Landscape in Argentina</td>
</tr>
<tr>
<td>I2d</td>
<td>Direct</td>
<td>Brown and white puppy</td>
</tr>
<tr>
<td>I2r</td>
<td>Recalled</td>
<td>Dog breed “dogos” in Argentina/Participant’s own pet</td>
</tr>
<tr>
<td>I3d</td>
<td>Direct</td>
<td>Girl sticking her tongue out</td>
</tr>
<tr>
<td>I3r</td>
<td>Recalled</td>
<td>A child in participant’s life/Children in Argentina</td>
</tr>
<tr>
<td>I4d</td>
<td>Direct</td>
<td>Storm over New York City skyline</td>
</tr>
<tr>
<td>I4r</td>
<td>Recalled</td>
<td>A city in Argentina (e.g. Buenos Aires)</td>
</tr>
<tr>
<td>I5d</td>
<td>Direct</td>
<td>Plate of Mexican food with guacamole and grasshoppers</td>
</tr>
<tr>
<td>I5r</td>
<td>Recalled</td>
<td>Unusual food in Argentina or that participant has tried</td>
</tr>
<tr>
<td>I6d</td>
<td>Direct</td>
<td>Traffic on highway overpass</td>
</tr>
<tr>
<td>I6r</td>
<td>Recalled</td>
<td>Traffic in Buenos Aires</td>
</tr>
<tr>
<td>I7d</td>
<td>Direct</td>
<td>Beach with hammock and palm trees at sunset</td>
</tr>
<tr>
<td>I7r</td>
<td>Recalled</td>
<td>Sunsets/beaches in Argentina</td>
</tr>
<tr>
<td>I8d</td>
<td>Direct</td>
<td>Strawberry cream cake</td>
</tr>
<tr>
<td>I8r</td>
<td>Recalled</td>
<td>Desserts in Argentina</td>
</tr>
</tbody>
</table>

Auditory stimuli were MP3 files that were played on an iPod Nano MP3 player (Model A1236). Participants wore headphones to listen to three separate tracks which they could replay as many
times as they wished. The first track featured the sound of birds chirping and had a playing time of 15 seconds (0:15). The second track lasted 14 seconds (0:14) and consisted of the sound of a blender motor in operation. The third track presented the sound of rain falling, and its duration was nine seconds (0:09). Audio files were obtained at no cost from the website freesfx.co.uk during the months of January and February, 2012. The set of contrasting direct evidence and recalled experience auditory stimuli appear in Table 4.6.

Table 4.6. Controlled Elicitation Task: Auditory Stimuli: Sounds versus Recalled Sounds

<table>
<thead>
<tr>
<th>Stimulus code</th>
<th>Stimulus type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1d</td>
<td>Direct</td>
<td>Birds chirping</td>
</tr>
<tr>
<td>S1r</td>
<td>Recalled</td>
<td>Insects buzzing, like flies, bees, or mosquitoes</td>
</tr>
<tr>
<td>S2d</td>
<td>Direct</td>
<td>Blender motor in operation</td>
</tr>
<tr>
<td>S2r</td>
<td>Recalled</td>
<td>Vacuum cleaner in operation</td>
</tr>
<tr>
<td>S3d</td>
<td>Direct</td>
<td>Rain falling</td>
</tr>
<tr>
<td>S3r</td>
<td>Recalled</td>
<td>Wind blowing</td>
</tr>
</tbody>
</table>

Olfactory stimuli consisted of essential oils diluted in distilled water and sprayed onto strips of perfume testing paper. The oils used in the task were selected under the guidance of an aroma therapist for their non-toxic, non-irritating properties. The oils (lavender, citrus, and a menthol/eucalyptus blend) were sourced from the brand Young Living Essential Oils. To create the sprays, the oils were mixed in quantities of 50 to 100 drops with approximately 50 millilitres of distilled water, and then stored in brown glass atomizer bottles. The set of contrasting direct evidence and recalled experience olfactory stimuli appear in Table 4.7.
Table 4.7. Controlled Elicitation Task: Olfactory Stimuli: Scents versus Recalled Scents

<table>
<thead>
<tr>
<th>Stimulus code</th>
<th>Stimulus type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1d</td>
<td>Direct</td>
<td>Lavender</td>
</tr>
<tr>
<td>A1r</td>
<td>Recalled</td>
<td>Rose or other flower</td>
</tr>
<tr>
<td>A2d</td>
<td>Direct</td>
<td>Menthol/eucalyptus blend</td>
</tr>
<tr>
<td>A2r</td>
<td>Recalled</td>
<td>Medicine (e.g. cough syrup)</td>
</tr>
<tr>
<td>A3d</td>
<td>Direct</td>
<td>Citrus</td>
</tr>
<tr>
<td>A3r</td>
<td>Recalled</td>
<td>Other fruit (e.g. apple or melon)</td>
</tr>
</tbody>
</table>

Tactile stimuli were three objects disguised in brown paper bags: a bundle of feathers, a sea sponge, and a palm-sized stone from a river in Canada. The stimuli were all personal possessions of the researcher of little or no monetary value. The set of contrasting direct evidence and recalled experience tactile stimuli appear in Table 4.8. Photographs of the tactile stimuli appear in Appendix D.
Table 4.8. Controlled Elicitation Task: Tactile Stimuli: Objects versus Recalled Textures

<table>
<thead>
<tr>
<th>Stimulus code</th>
<th>Stimulus type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1d</td>
<td>Direct</td>
<td>Bundle of feathers</td>
</tr>
<tr>
<td>O1r</td>
<td>Recalled</td>
<td>Hair of a pet (e.g. a cat or a dog)</td>
</tr>
<tr>
<td>O2d</td>
<td>Direct</td>
<td>Sea sponge</td>
</tr>
<tr>
<td>O2r</td>
<td>Recalled</td>
<td>Sandpaper</td>
</tr>
<tr>
<td>O3d</td>
<td>Direct</td>
<td>Stone</td>
</tr>
<tr>
<td>O3r</td>
<td>Recalled</td>
<td>Wood</td>
</tr>
</tbody>
</table>

Gustatory stimuli consisted of common, non-allergenic food items that participants were familiar with: potato chips, grapes, and caramel candies. The stimuli were purchased at supermarkets in Buenos Aires at the time of testing. The set of contrasting direct evidence and recalled experience gustatory stimuli appear in Table 4.9.
Table 4.9. Controlled Elicitation Task: Gustatory Stimuli: Foods versus Recalled Tastes

<table>
<thead>
<tr>
<th>Stimulus code</th>
<th>Stimulus type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1d</td>
<td>Direct</td>
<td>Potato chips</td>
</tr>
<tr>
<td>T1r</td>
<td>Recalled</td>
<td>Nuts</td>
</tr>
<tr>
<td>T2d</td>
<td>Direct</td>
<td>Grapes</td>
</tr>
<tr>
<td>T2r</td>
<td>Recalled</td>
<td>Raisins</td>
</tr>
<tr>
<td>T3d</td>
<td>Direct</td>
<td>Caramel candies</td>
</tr>
<tr>
<td>T3r</td>
<td>Recalled</td>
<td>Chocolate</td>
</tr>
</tbody>
</table>

4.2.2.3 Procedure

I explained the instructions prior to the start of the task. Participants heard that they would be presented with a series of items for each of the five senses, and that there would be eight pictures for sight, three sounds for hearing, three aromas for smell, three objects for touch, and three food items for taste. I asked participants to describe their sensory impressions of the items in their own words, using one or more sentences. Participants were then shown the practice image, and I provided several examples of possible responses, none of which included the target structure of ser or estar with an adjectival predicate. I confirmed understanding of the instructions with the participants and checked whether participants had any allergies or sensitivities before proceeding with the recording. Once recording was underway, I presented the stimuli in the order described above, where a direct sensory stimulus was immediately followed by a question contrasting the direct stimulus with a recalled experience involving the same sense.
4.2.3 Guided Interview Task

4.2.3.1 Background

The purpose of the interview was to record a sample of vernacular speech from each participant. Taken together, the interviews made up a corpus of natural language from which spontaneous examples of clauses that contained *ser* or *estar* in combination with an adjectival predicate were later extracted and analyzed in their linguistic context. The data was also analyzed according to the social variables of age and level of education in order to evaluate what role, if any, they play in the selection of *ser* and *estar*.

4.2.3.2 Method

The guided sociolinguistic interview was the least constrained of the three tasks. Because an interview is essentially a conversation between two people, much depends on the interpersonal dynamics between the interviewer and the interviewee. The length and nature of the conversation can vary tremendously from one interview to the next. However, in the interests of establishing the optimal conditions for consistent and successful interviews, I followed a standard sociolinguistic interview script that was adapted from Tagliamonte’s (2006) *Interview Schedule: Guideline Questions (Adapted from Labov, 1973)*. Several modifications were made to Tagliamonte’s set of questions. The questions were translated into Buenos Aires Spanish, and I adjusted the vocabulary for specific social and cultural references (e.g. holidays, food, education). A limited number of questions were selected with the objective of stimulating maximum production of *ser* and *estar* with adjectival predicates within an estimated 15-minute time frame for the interview. None of the questions contained the structure that is the focus of the study, to avoid any potential priming effects. The complete script appears in Appendix E.
4.2.3.3 Procedure

All tasks in the study were conducted in one session in a location chosen for the participant’s convenience. Locations included the participant’s residence, the researcher’s residence, the participant’s place of employment, various coffee shops, and in one case a public park. After participants had completed the language and personal background questionnaire, I began recording the oral tasks (the guided interview and the controlled elicitation tasks). A Phillips Voice Tracker digital audio recorder, Model LFH0645, was used to record MP3 files of these two tasks. In all but three cases, the oral tasks were recorded in a single MP3 file. In the remaining three cases, due to an interruption, there are two MP3 files that constitute the participant’s oral tasks. The duration of the interviews ranged from 5:00 (five minutes) to 32:35 (32 minutes and 35 seconds). The mean duration of the interview was 16:08 (16 minutes and eight seconds).

4.3 Analysis

The data were transcribed (for the two oral tasks), coded, and analyzed both quantitatively and qualitatively. I performed the transcription, coding, and where relevant, translation of the data. In any cases where there was doubt about a transcription, I consulted a native speaker of Buenos Aires Spanish for confirmation of the transcription. The software I used for transcription was Audacity (version 1.2.5; Mazzoni & Dannenberg, 2000).

All data were coded using ser/estar selection as the dependent variable. The social variables of age and level of education were included in the coding for all tasks, and other independent variables differed according to the task. The adjectives in the data from the oral tasks were coded according to semantic type, and categorical versus variable [copula + adjective] contexts were identified. The quantitative analysis was based on variable contexts only, whereas the qualitative analysis evaluated both categorical and variable contexts.
The quantitative analysis comprised a distributional analysis and a multivariate analysis for each task. The factor-by-factor (i.e., distributional) analysis examined each of the factors separately. I carried out the factor-by-factor analysis in Microsoft Excel (version 15.30) and R (R Core Team, 2016). It tested how each factor, taken in isolation, influenced the use of *ser* and *estar* with adjectives. This provided some preliminary insights into the conditioning factors on the \[ser/estar + \text{adjective}\] context. The multivariate analysis then provided a view of how these constraints combined to produce the overall distributional picture of the selection of *ser* and *estar* with adjectives.

For the multivariate analysis, I conducted a mixed-effects logistic regression analysis in R (version 3.3.1; R Core Team, 2016), using the Rbrul package (Johnson, 2009). I chose to use Rbrul because it offered the advantages of a mixed model. A mixed model includes both fixed effect and random effect factors (Johnson, 2009, p. 364). Random effect factors have a large number of levels, which are randomly drawn from a wide selection of possibilities. By contrast, fixed effect factors have a small number of levels, which are limited. Fixed effect factors in the data included several of the independent variables such as age, level of education and presence/absence of evidence in the context (others varied according to the task). Speaker was a random effect factor in the data for all three tasks. Johnson (2009) explains that “including a speaker random effect takes into account that some individuals might favour a linguistic outcome while others might disfavour it, over and above (or ‘under and below’) what their gender, age, social class, etc. would predict” (p. 365). The other random effect factors included Item in the CPT and Adjective in the two oral tasks.

The results are reported following the conventions in the field of language variation and change, where statistical analysis has often been conducted using GoldVarb (Sankoff et al., 2005). The format and terminology used in reporting the results follows these conventions, but I anticipate that they may be unfamiliar to those in other fields of study. For this reason, I have included explanatory notes along with the tables reporting the results of the multivariate analyses in the next chapter.
5.0. Introduction

For reasons outlined in the description of the methodology in the previous chapter (see section 4.2.1.4. for discussion), the order in which the participants completed the tasks was the reverse of the order in which I present and describe the tasks. In this chapter, I report and discuss the results of each of the three tasks, starting with the CPT, followed by the elicitation task, and finally the interview task. I consider the relevance of my findings to the broader field of variation and change in the [ser/estar + adjective] context, and evaluate the contribution of these data to theories of copulas and adjectives.

5.1 Contextualized Preference Task

To recap, in the contextualized preference task, participants read an introductory context and were then prompted to choose between sentences with either ser or estar and an adjectival predicate, as in (106).

(106)  

A. Bien, el patio es grande.

‘A. Well, the patio is SER big.’

B. Bien, el patio está grande.

‘B. Well, the patio is ESTAR big.’

The task also included a second opportunity to choose between ser and estar in the event that the participant found both verbs to be appropriate for the context (107).
If both sentences are acceptable, and you were going to say this, how would you say it?

___ The first ___ The second ___ I can’t decide.’

In most cases, these two layers of questioning forced a choice, which was the goal of the task. In a few cases, however, participants maintained that both verbs were acceptable, and selected No puedo decidir, ‘I can’t decide’ at the second stage. The fact that some participants chose both verbs in some contexts is to be expected and is consistent with my hypothesis. Based on this reasoning, the both contexts are included in the estar counts in the quantitative analysis, and in the text, I refer to the combined estar + both counts as one category of estar-compatible responses.

5.1.2 Evidential versus neutral contexts

The first analysis takes a global look at whether estar was selected at a higher frequency in contexts where evidence was present than in contexts where evidence was not present. I predicted that speakers would show a difference in patterns of selection of ser and estar in their responses depending on whether or not evidence was present in the context. The results in Table 5.1 show that speakers selected estar (or both verbs) more frequently in contexts where evidence was present than in the neutral contexts.
Table 5.1. Frequency of selection of *ser* and *estar:both according to presence of evidence in the context

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Ser</th>
<th>Estar/Both</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>N</td>
<td>413</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>76.5</td>
<td>23.5</td>
</tr>
<tr>
<td>Neutral</td>
<td>N</td>
<td>490</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>90.7</td>
<td>9.3</td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td>903</td>
<td>177</td>
</tr>
</tbody>
</table>

The proportion of the selection of *estar*-compatible responses in contexts where evidence was present was greater than the proportion of the selection of *estar*-compatible responses in contexts where evidence was not present. The frequency of responses that included *estar* (*estar* or both verbs selected) in evidence-present versus evidence-neutral contexts was submitted to a chi-square test in the statistical package R (R Core Team, 2016). A highly positive correlation was found between the presence of evidence in the context and the selection of *estar* or both verbs ($\chi^2 = 39.03, df = 1, p < .001$).

5.1.3 Type of evidence

The second analysis is a more detailed evaluation of each of the three types of evidential contexts (direct, inferred, and reported evidence). This analysis considers whether speakers showed a difference in frequency of selection of *estar* according to the type of evidence present in the context. My prediction for this variable was based on the proposed implicational hierarchies for classifying types of evidence across languages (Aikhenvald, 2004; Speas, 2004; Willett, 1988). If a language specifies only one type of evidence, then it is direct evidence. I therefore expected
that contexts of direct evidence would show a higher frequency of *estar* use than contexts of inferred or reported evidence. The results in Table 5.2 show that this prediction was partially borne out. Speakers selected *estar* more frequently in contexts where either direct or reported evidence was present than in contexts where inferred evidence was present.

Table 5.2. Frequency of selection of *ser* and *estar*/both by evidence type in evidence-present versus evidence-neutral contexts

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Present</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ser</td>
<td><em>Estar</em>/both</td>
</tr>
<tr>
<td>Direct</td>
<td>N 122</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>% 67.8</td>
<td>32.2</td>
</tr>
<tr>
<td>Inferred</td>
<td>N 171</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>% 95</td>
<td>5</td>
</tr>
<tr>
<td>Reported</td>
<td>N 120</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>% 66.7</td>
<td>33.3</td>
</tr>
<tr>
<td>Total</td>
<td>N 413</td>
<td>127</td>
</tr>
</tbody>
</table>

The frequency of selection of *estar*-compatible responses varied according to evidence type. For contexts where direct evidence was present, we observe the greatest impact of evidence on the selection of *ser* and *estar*. In the direct evidence condition, the frequency of selection of responses that included *estar* in the evidential context (58) was approximately eight times that of the frequency of selection of responses that included *estar* in the neutral context (7); in other words, there was a ratio of 8:1 in the selection of *estar*-compatible responses in the direct evidence context over the neutral context. In the reported evidence condition, there was also an effect of evidence, albeit to a lesser degree, reflected in a ratio of about 2:1 in the selection of
estar-compatible responses in the reported evidence context over the neutral context. The frequency of selection of responses that included estar in contexts where reported evidence was present (60) was about twice that of neutral contexts (29). By contrast, in the inferred evidence condition, there was no significant contribution of evidence as measured by a difference in frequency of selection of ser and estar. The frequency of selection of estar-compatible responses in contexts where inferred evidence was present (9) was slightly less than the frequency of selection of responses that included estar in neutral contexts (14).

The frequencies of responses classified as to whether they included estar or not (estar or both verbs selected) for each pair of evidence-present and corresponding neutral conditions were submitted to chi-square tests in R. For the direct versus neutral condition, the result was highly significant ($\chi^2=46.94$, $df=1$, $p<.001$). For the reported versus neutral condition, the result was again highly significant ($\chi^2=13.43$, $df=1$, $p<.001$). For the inferred versus neutral condition, the result was not significant ($\chi^2=0.74$, $df=1$, $p=.388$).

5.1.4 Social variables

The third analysis considers the roles that the extralinguistic variables of age and level of formal education played in the use of estar as an evidentiality strategy. This analysis examines whether speakers of different age groups (younger vs. older) and educational backgrounds (less vs. more formal education) selected estar at higher rates in contexts where evidence was present than in contexts where evidence was not present. Based on findings of previous studies (Cortés-Torres, 2004; Gutiérrez, 1992; Silva-Corvalán, 1986), I predicted that younger speakers and less educated speakers would show greater rates of selection of estar in contexts where evidence was present than in contexts where evidence was not present as compared with older speakers and more educated speakers. The results for the frequency of selection of ser and estar in the two types of contexts according to social variables are shown in Table 5.3 (age) and Table 5.4 (level of education).
Table 5.3. Frequency of selection of *ser* and *estar* / both in evidence-present versus evidence-neutral contexts according to speaker age

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Younger (age 18-30)</th>
<th>Older (age 50+)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Ser</em></td>
<td><em>Estar</em> / both</td>
</tr>
<tr>
<td>Present</td>
<td>N</td>
<td>227</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>75.7</td>
</tr>
<tr>
<td>Neutral</td>
<td>N</td>
<td>273</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>91</td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td>500</td>
</tr>
</tbody>
</table>

It is evident from the results in Table 5.3 that there was very little contrast between the behaviour of younger and older speakers. Therefore, speaker age does not appear to be a significant factor in the selection of *estar*-compatible responses in contexts where evidence was present as compared with contexts where evidence was not present. This result runs counter to my prediction and confirms the null hypothesis.
Table 5.4. Frequency of selection of *ser* and *estar*/both in evidence-present versus evidence-neutral contexts according to speaker level of education

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Less formal education</th>
<th>More formal education</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(secondary or less)</td>
<td>(at least some post-secondary)</td>
</tr>
<tr>
<td></td>
<td><em>Ser</em></td>
<td><em>Estar</em> / both</td>
</tr>
<tr>
<td>Present</td>
<td>N</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>75</td>
</tr>
<tr>
<td>Neutral</td>
<td>N</td>
<td>222</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>88.1</td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td>411</td>
</tr>
</tbody>
</table>

Based on the results in Table 5.4, it appears that less educated speakers were slightly more inclined than more educated speakers to choose *estar* or both verbs in either type of context. This finding partially confirms the prediction that less educated speakers would show greater rates of selection of *estar* in contexts where evidence was present than in contexts where evidence was not present as compared with more educated speakers. Less educated speakers did show greater rates of selection of *estar* than more educated speakers. However, this was true not only in contexts where evidence was present, but also in contexts where evidence was not present.

5.1.5 Multivariate analysis

The multivariate analysis included the three fixed-effect predictors Evidence Type, Age and Education. The original six levels of the predictor Evidence Type, containing both evidential and neutral conditions for each of the categories of the tripartite division of direct/inferred/reported evidence, were reduced to five after finding that the model with five levels was a better fit. This
decision was based on the results of the factor-by-factor analysis, where I found that rates of
estar selection in direct and reported evidence contexts were comparable, and were significantly
different from those in inferred evidence contexts. Thus, the two categories of direct and reported
evidence were collapsed into a single level of the predictor Evidence Type. The two direct and
reported categories for the corresponding neutral contexts formed two separate levels, and the
inferred evidential and neutral contexts made up the other two levels of this predictor. I tested for
all possible interactions between fixed effects in the same model (Evidence Type: Age, Evidence
Type: Education, and Age: Education). Two random-effect predictors, Speaker and Item, were
included. The results reported in Table 5.5 were taken from the best stepping-down iteration of
the logistic regression.

A brief explanatory note is in order for readers who may be unfamiliar with the format and
terminology of the tables used to report the results of logistic regression analyses in this chapter.
This method of reporting factor effects is based on sum contrasts, and reports factor weights for
fixed effects (probabilities ranging from 0 – 1), which are coefficients that represent deviations
from the mean. A factor weight of 0.5 indicates no effect. A method that reports logodds uses
treatment contrasts, taking one level of each factor as a baseline (with a coefficient of 0) and
reporting a coefficient for each additional level. The value of the logodds is stated as a positive
or negative value, and corresponds to the probability of the effect. A factor weight greater than
0.5 is equivalent to a logodds with a positive value, and these values indicate a favouring effect.
A factor weight below 0.5 is equivalent to a logodds with a negative value, and these values
indicate a disfavouring effect. (Johnson, 2009, p. 361). For convenience, I report both factor
weights and logodds in the multivariate analysis tables.

9 The values reported for the non-significant fixed effects were taken from the full model (the first model in the
step-down analysis).
Table 5.5. Mixed-effects logistic regression analysis of the contribution of external and internal factors to the probability of *estar* selection in the contextualized preference task with Speaker and Item as random effects.

<table>
<thead>
<tr>
<th>Input</th>
<th>Total N</th>
<th>0.064</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Significant Fixed Effects: |         |       |
| EVIDENCE TYPE              |         |       |
| Direct + Reported (Evidential) | 0.849  | 1.728 | 32.8 | 360 |
| Reported (Neutral)          | 0.642  | 0.586 | 16.1 | 180 |
| Inferred (Neutral)          | 0.48   | -0.079| 7.8  | 180 |
| Inferred (Evidential)       | 0.298  | -0.857| 5    | 180 |
| Direct (Neutral)            | 0.201  | -1.378| 3.9  | 180 |
| Range^b                     | 64.8   |       |      |     |

| Non-significant Fixed Effects: |         |       |
| AGE                          |         |       |
| Younger                      | [0.511]c | 0.045 | 16.7 | 600 |
| Older                        | [0.489] | -0.045| 16   | 480 |

| EDUCATION                   |         |       |
| Less                         | [0.557] | 0.23  | 18.5 | 504 |
| More                         | [0.443] | -0.23 | 14.6 | 576 |

| Random Effects:^d            |         |       |
| SPEAKER                      | sd = 0.908 |       | N = 45 |
| ITEM                         | sd = 1.532 |       | N = 24 |

a. FW: Factor Weight.
b. The value of the range is determined by calculating the difference between the values of the factor weights of the highest and lowest ranked factor levels within a factor group. A comparison of the value of the range for different factor groups establishes a ranking of significance across factor groups. Factor groups with the greatest range are most significant and those with the smallest range are least significant (Tagliamonte, 2006, p. 235).
c. Factor weights for non-significant fixed effects are distinguished by square brackets.
d. The model produces a single estimate of the amount of variation for random effect factors (Johnson, 2009, p. 365).
The predictor Evidence Type was significant \( (p = 0.0433) \), but Age and Level of Education were not significant predictors of \textit{estar} selection in the model. No interactions between fixed effects were significant \( (p\text{-values were} > 0.05) \). The constraint ranking for Evidence Type revealed that the direct and reported evidential contexts strongly favoured \textit{estar} (0.849). This result confirmed the prediction, based on the findings of the factor-by-factor analysis, that direct and reported evidence contexts would show evidential marking. A surprising result was that reported neutral contexts favoured \textit{estar} (0.642). The higher ranking of the neutral contexts (0.48) as compared with evidential contexts (0.298) for inferred evidence was another unexpected result, but neither category favoured \textit{estar}, as predicted. Nor was \textit{estar} favoured in the category of direct neutral contexts (0.201), which was similarly unsurprising.

Based on these results, we can reject the null hypothesis, and conclude that \textit{estar} is selected to mark the presence of direct and reported evidence in the context. As for the social variables tested, there was no support for either of the predictions with respect to age and level of education. Neither younger nor less-educated speakers chose \textit{estar} at significantly higher frequencies than their older or more-educated counterparts. Although these predictors were not significant, the ranking of the constraints for Age and Education matched the predictions.

5.1.6 Summary and discussion: Contextualized Preference Task

In sum, speakers selected \textit{estar} at significantly greater frequencies when direct or reported evidence was present in the context than in neutral contexts or in contexts of inferred evidence in the contextualized preference task. The finding that \textit{estar} selection was associated with direct and reported evidence, but not inferred evidence, was initially unexpected. However, what distinguishes inferred evidence from direct and reported evidence is that it is the site of overlap between epistemic and evidential meaning (Faller, 2002). A possible explanation for this result may be found in the existence of a range of alternative means of encoding epistemic meaning in Spanish. The subjunctive in Spanish is a robust grammatical paradigm for the expression of epistemic meaning, and epistemic modals represent a further means of expression. An additional
option for the expression of inferred evidence in Spanish is the synthetic future tense, which has developed an extension of conjectural (or inferred) meaning, as in (108).


   ‘The phone rings. “It must be my sister,” says Juana.’

Because epistemic meaning is already associated with a variety of other forms, it does not seek expression as an extension of *estar*. By contrast, direct and reported evidential meanings have no grammatical means of expression in Spanish, hence their association with *estar*. Cross-linguistically, it is not unusual for different types of evidential meaning to be realized via diverse morphosyntax\(^1\) (see section 2.6.2 for discussion). Based on this analysis, we can conclude that in Spanish, there is a difference between inferred evidence, which is realized through various other forms in the language (or is simply not marked), and direct and reported evidence, which are semantically encoded in the extension of *estar*.

With respect to social variables, similar patterns of contrasting rates of *estar* selection occurred across speakers of different ages and levels of education in both evidential and neutral contexts. These findings indicate that the evidential use of *estar* in Buenos Aires Spanish reflects a situation of stable variation, rather than a change in progress. In the next section, I evaluate whether the results of the elicitation task reinforce or repudiate these conclusions.

### 5.2 Controlled Elicitation Task

There were two stages in the elicitation process for each stimulus in the controlled elicitation task. The stimulus for direct evidence was presented first, and then participants were prompted to recall and describe something similar from their own previous experience. These are referred to in this section as direct evidence and recalled experience contexts, respectively.

---

\(^1\) Blain and Dechaine’s (2007) proposal of an Evidential Domain Hypothesis offers an account of how evidentials can be introduced into the clause in several different positions, depending on the type of viewpoint that the speaker expresses.
5.2.1 Coding

In what follows, I describe the approach to coding the data and some procedural decisions regarding the inclusion of an utterance in the quantitative analysis. For the coding, in cases of coordination, adjectives were counted as separate utterances with the preceding verb if there was an overt coordinator, for example, \(y/e\) ‘and’, \(o/u\) ‘or’, \(pero\) ‘but’, as in (109).

(109) \(… \)como que estaban todos muy, muy \textit{apretados e incómodos} [la vivienda de los inmigrantes en el pasado]…

‘… like they were \textit{ESTAR} all very, very \textbf{cramped and uncomfortable} [the living situation of immigrants in the past]…’

(CM0417:I4r)

If the speech contained a pause, transcribed as a comma, the adjective(s) following the pause were not counted as a complete [verb + adjective] utterance (110).

(110) \(… \)bueno, está muy pulida, muy gastada…[la piedra]…

‘…well, it \textit{ESTAR} very polished, very washed-out… [the stone]…’

(SM0426:O3d)

Repetitions, such as those in (111), were included in the count.

(111) \(… \)\textit{está bueno} para hacer catarsis, par, sí, para el estrés, sí, \textit{está bueno}…

‘… \textit{it \textit{ESTAR} good} for experiencing catharsis, for, yes, for stress, yes, \textit{it \textit{ESTAR} good}…’

(EG0403:O2d)
5.2.1.1 Coding of verbs

Initially, all responses that contained an adjective were transcribed, regardless of whether or not *ser* or *estar* appeared in the utterance. In three sample responses (112-114) to a stimulus for direct visual evidence\(^{11}\), there were five phrases that contained adjectives. The two utterances that contained *ser* or *estar* were counted in the quantitative analysis, and the other three were excluded.

(112) *...es hermosa. Sí, muy alegre. Se ve que está feliz.*

‘...she’s *SER* pretty. Yes, very joyful. It shows that she *ESTAR* happy.’

(EG0403:I3d)

(113) *...también me parece tierna...*

‘...also she looks sweet to me ...’

(CP0331:I3d)

(114) *...ay, qué bonita...*

‘...aww, how pretty...’

(GS0418:I3d)

In (112), there were three occurrences of an adjective in the participant’s response: the adjective occurred with *ser* in the first phrase, the verb was omitted in the second utterance, and *estar* appeared in the third phrase. The response in (113) contained the semi-copula verb *parecer*, ‘to appear’, with the dative pronoun *me*, ‘to me’. The speaker employed the exclamative *qué*, ‘what/how’ with an adjective in the utterance in (114). Arguably, all of these structures convey different meanings, and emphasize different semantic, pragmatic and stylistic choices. A

\(^{11}\) The stimulus for these utterances was Image 3, in which a smiling young girl makes a playful gesture (see Appendix D).
function-based analysis would consider utterances that included verb omission, semi-copula verbs, exclamative phrases, as well as other structures such as presentational *haber* (‘there is/there are’). However, I conducted a form-based analysis here in order to test the hypothesis that a speaker’s choice to use *ser* or *est*ar with an adjectival complement expressed a difference in evidential meaning. My goal was to find evidence to support the existence of this contrast, and as a result, the data was restricted to only those utterances that contained the [*ser/estar + adjective*] structure. The two category labels in the coding of verbs were *ser* and *estar*.

5.2.1.2 Coding of adjectives

The masculine singular form of the adjective was used for coding, and all affixes were stripped from the adjective. Examples of the types of morphemes that were removed are emphatic prefixes and suffixes (*hiper-*-, *re-*-, *contra-*-, *-ísimo*), and diminutive and augmentative affixes (*–ito*, *–ón*, *–ote*). Adjectives were coded according to a variety of syntactic and semantic criteria. I begin with the adjective types that categorically appeared with *ser* in the data, and which were excluded from the counts (there were no adjective types that appeared exclusively with *estar*). I then describe the variable adjective types, several of which were also excluded from the quantitative analysis for reasons which will become clear. A qualitative analysis follows the results of the quantitative analysis later in this chapter. There, I consider the behaviour of some types of adjectives in the data and discuss what these observations can contribute to our understanding of the evidential use of *estar*.

There were five types of adjectives that appeared categorically with *ser* in the data. They comprise intensional, dispositional, relational and kind-level adjectives, as well as adjectives with adverbial readings. The semantics of these categories of adjectives are distinct from those of property-denoting adjectives, for diverse reasons.

---

12 Given that predicative adjectives are always intersective (Hofherr, 2010, p. 7), I did not include in the coding any of the different types of non-intersective adjectives that may appear in attributive position.
(i) Intensional adjectives

Demonte (2008) analyzes the predicative use of intensional adjectives as implicit relative clauses in which the adjectives have a qualificational sense. These types of adjectives include modal/epistemic adjectives (e.g. posible ‘possible’, necesario ‘necessary’, falso ‘false’), and degree/quantifying adjectives (e.g. completo, ‘complete’, rotundo ‘total’, simple ‘simple’, puro ‘pure’, mero ‘mere’, exclusivo ‘exclusive’, único ‘only’) (Demonte, 2008, pp. 82-83). In this data, there were 10 utterances with modal/epistemic adjectives and five utterances with degree/quantifying adjectives. The sentences in (115) and (116) are examples from the data of modal/epistemic and degree/quantifying adjectival predicates, respectively.

(115) ...la lluvia es necesaria para el campo...

‘...rain is\textit{ necessary for the countryside}...’

(EP0404:S3d)

(116) ...ese Kaiser [el perro] era único...

‘...that Kaiser [the dog] \textit{was} unique...’

(CC0403:I2r)

(ii) Dispositional adjectives

Dispositional adjectives (also called eventive, dynamic or behavioural adjectives) include such adjectives as cruel ‘cruel’, tonto ‘stupid’, cortés ‘polite’, and discreto ‘discreet’. These adjectives are individual-level predicates that give rise to a behavioural reading in certain contexts with \textit{ser}, such as the present progressive, the imperative and with a prepositional complement (117). Various authors have analyzed them as combining with an event argument at some level (Arche, 2006; Fábregas et al., 2012\textsuperscript{13}; Marín, 2010; Schmitt, 2005).

\textsuperscript{13} Fábregas et al. (2012) use the term \textit{evaluative} to classify these adjectives. I avoid this usage in order to remain consistent with the use of \textit{evaluative} throughout the current study (see section 6.3.3.).
Juan fue muy cruel con Pedro.

‘Juan was very cruel with Pedro.’

(Arche, 2006, p. 92)

Only one utterance with a dispositional adjective appeared in this data, with ser (118).

...han hecho de Nueva York una ciudad que hoy por hoy hasta, a las tres, cuatro de la mañana podés volver al hotel caminando, está limpia, la gente sigue siendo loca, original, cosmopolita...

‘…they have made New York a city that, for the time being, until three, four in the morning you can go back to the hotel on foot, it is clean, the people keep on acting/being crazy, original, cosmopolitan…’

(CN0405:14r)

(iii) Relational adjectives

Relational adjectives are often denominal and encompass such items as político ‘political’, eléctrico ‘electric’, and vegetariano ‘vegetarian’. McNally and Boleda (2004) argue against the traditional semantic analysis of relational adjectives as predicate modifiers and propose instead that relational adjectives denote properties of kinds. As such, they contain a kind-level argument. These authors suggest that relational adjectives and kind-level adjectives (Krifka et al., 1995) are of the same semantic type. McNally and Boleda’s analysis and the semantic similarity of relational adjectives to nouns lead us to predict that relational adjectives should appear with ser in Spanish. However, Escandell-Vidal and Leonetti (2002) point out that many individual-level predicates, including relational adjectives, can be coerced into evidential readings with estar. Nonetheless, relational adjectives occurred exclusively with ser in this data (N=19), as shown in (119).
...y después mi, mi abuela paterna hacía unos postres pero **eran italianos**, eran, no eran com-, no eran de acá…

‘…and then my, my paternal grandmother used to make some desserts but **they were** Italian, they were, they weren’t like-, they weren’t from here…’

(MP0322:18r)

(iv) Kind-level adjectives

A kind-level adjective denotes a property of a kind, and not of an individual, hence their combination with *ser* in Spanish. Krifka et al. (1995) identify the class of kind-level adjectives in English as consisting of *extinct, widespread, scarce, abundant, common* and *rare*. There were a total of 10 utterances in the data with the two kind-level adjectives común ‘common’ and *raro* ‘rare’. An example is given in (120).

(120) ...*acá es común, el canto de los pájaros*...

‘…here it is **common**, the singing of birds (birdsong)…’

(MT0407:S1d)

(v) Adjectives with adverbial readings

There are several classes of adjectives with adverbial readings, including frequency adjectives such as *occasional* and *frequent*, as well as the adjectives *same* and *different* (see Morzycki, 2016, pp. 59-78, for analysis and discussion). These types of adjectives tend to appear in attributive position, but some can appear in predicative position. They categorically combined with *ser* here. Frequency adjectives in the data included *habitual* ‘usual’, *usual* ‘normal’, *típico* ‘typical’, *característico* ‘characteristic’, *constante* ‘constant’, and *popular* ‘popular’. The adjectives that corresponded to *same* and *different* here are *diferente* ‘different’, *distinto*
‘different’, similar ‘similar’, igual ‘same’, desigual ‘unequal’, and parecido ‘similar’. There were 46 utterances with frequency adjectives and 14 utterances with same/different adjectives in the data. The sentences in (121) and (122) are example utterances that contain frequency and same/different adjectival predicates, respectively.

(121) …es un sonido que tal vez para escuchar un rato no es molesto pero si es constante, sí…
‘…it is a sound that maybe to listen to for a while is not annoying but if it is constant, yes …’

(CP0331:S2d)

(122) …[los gatos] son muy distintos a los perros, obviamente…
‘…[cats] are very different from dogs, obviously…’

(NM0405:I2r)

Five types of adjectives were variable in their combination with ser and estar in the data. These included locative predicates, polysemous, perfective and qualificational adjectives, as well as adjectival participles.

(vi) Locative predicates

There were nine uses of ser and estar with the locative predicates cerca ‘near’ and lejos ‘far’ in the data¹⁴. When the subject is an individual, locative predicates are predicted to select estar, in an expression of the original lexical meaning of spatial location of this verb (123). Eventive subjects are expected to appear with ser, but no utterances contained eventive subjects in this data.

(123) …cuando estás cerca de un aeropuerto, los ruidos…

¹⁴ Cerca and lejos are classified as adverbs, not adjectives, in the dictionary of the Real Academia Española.
‘...when you are \textit{ESTAR} near an airport, the noises...’

(ES0411:S2r)

There was variability between \textit{ser} and \textit{estar} across locative predicates with subjects that denote locations in this data (124-125).

(124) \textit{[Pinamar] está más cerca que Mar del Plata...}

‘[Pinamar] \textit{ESTAR} closer than Mar del Plata...’

(MC0327:I7r)

(125) \textit{Mar del Plata es lejos, 400 kilómetros...}

‘...\textit{SER} far, 400 kilometres...’

(ARG0410:I7r)

None of the responses with locative predicates were counted in the quantitative analysis, given that an evidential reading is not available when the use of \textit{estar} is determined by the lexical semantics of the verb.

(vii) Polysemous adjectives

Another exclusion on the basis of variation that precludes the possibility of evidential meaning was polysemous adjectives. The distinct lexical entries of polysemous adjectives are determined by the verb with which they combine (126). Such adjectives include \textit{listo} ‘clever’ (\textit{ser})/‘ready’ (\textit{estar}), and \textit{vivo} ‘sharp’ (\textit{ser})/‘alive’ (\textit{estar})

(126) i. \textit{Javier es listo.}

‘Javier \textit{SER} clever.’

ii. \textit{Javier está listo.}
‘Javier is \text{\textsc{ESTAR}} ready.’

The data contained nine adjectives that were polysemous, some of which showed the predictable, lexically determined variation with \textit{ser} and \textit{estar} as in (127-128).

(127) ...\textit{yo \textsc{estaba} débil y me, y me hicieron tomar un jarabe que se llama, eh, ah, pero es horrible, la cosa de que tengo recuerdo...} T: Aceite... MC: Aceite de hígado de bacalao.

\textit{Aceite de hígado de bacalao. Ahí está. Horroroso.}

‘...I \textit{was\textsc{ESTAR} sick} and they, and they made me drink a syrup that is called, uh, oh, but it \textit{is\textsc{SER} horrible, the thing that I remember...} T(Researcher): Oil... MC(Participant): Cod liver oil. Cod liver oil. That’\textit{SESTAR it. Horrific.}’

(MC0327:A2r)

(128) ...\textit{es muy débil.}

‘... it \textit{is\textsc{SER} very weak.}’ [describing the stimulus of a scent]

(EP0404:A1d)

(viii) Adjectival participles

When a participle is used as a resultative adjective in Spanish (129), \textit{estar} is syntactically selected to express this meaning (Bosque, 1990).

(129) ...\textit{hay algunas, algunas cosas que si están muy cocidas son, como, duras, ¿viste?}

‘... there are some, some things that \textit{if they are\textsc{ESTAR} very cooked} they are\textit{SER}, like, hard, you know?’

(BB0425:I5r)
Some participles combine with *ser* to denote individual-level predicates in which the resultative meaning is absent (130).

(130) ...*ella es más reservada, la nena*...

‘...she is more reserved, the little girl...’

(CC0419:13r)

Because the appearance of adjectival participles with *estar* is determined by the syntax-semantics, the possibility of identifying an evidential reading is ruled out. Hence, utterances that contained adjectival participles were excluded from the counts\(^\text{15}\) (N=149).

(ix) Perfective adjectives

Much like adjectival participles, perfective adjectives (or truncated participles) give rise to a resultative meaning (Bosque, 1990). These adjectives are also thought to be deverbal, and include items such as *lleno* ‘full’, and *suelto* ‘loose’. Again, they are predicted to combine with *estar* for syntactic reasons (although some perfective adjectives do appear with *ser* with an atemporal reading, for example, *limpio* ‘clean’, and *seco* ‘dry’). An evidential use of *estar* with a perfective adjective is indistinguishable from a resultative meaning derived from the syntactic-semantic composition, as in (131), where the participant responded to the direct evidence stimulus of an image of a hammock on a beach.

(131) ...*muy bonita. La, la hamaca está vacía, ¿no?*

‘...very beautiful. The, the hammock is empty, no?’

\(^{15}\) Utterances with verbal participles, in which a participle appeared with *ser* in a passive construction (i), were excluded from the data.

(i) ...*con algunos perros es como que son un poquitito anti-naturales, como que son muy, muy fabricados por el hombre, ¿viste?*

‘...with some dogs it is a tiny bit unnatural, like they are very, very fabricated by mankind, you know?’

(BB0425:12d)
Furthermore, an analysis of the selection of *ser* and *estar* with perfective adjectives must take into account the part-structure of the subject in order to account for the expression of quantization in the sentence. Because of the low number of tokens (N=87), an analysis of the utterances with perfective adjectives in the data would have been insufficient for statistical validity after coding for the various elements involved in the composition of meaning in these sentences. The data with perfective adjectives were therefore eliminated from the quantitative analysis (but for a qualitative analysis of this data, see section 5.2.3.1).

(x) Qualificational adjectives

The quantitative analysis was performed upon those utterances that contained qualificational adjectives, such as *bueno* ‘good’, *fuerte* ‘strong’, *picante* ‘spicy’, *suave* ‘soft’. These adjectives have a property-denoting semantics (Demonte, 2008) and were the most frequent of all adjective types in the data (N=1124). There was variation between *ser* and *estar* with many qualificational adjectives (132-133), and the quantitative analysis tested whether this variation can be attributed to the evidential use of *estar*.

(132)  ...*es muy lindo el cítrico*...

‘... *it isSER very pretty the citrus*...’

(EJB0421: A3d)

(133)  ...*está linda la foto*...

‘... *it isESTAR pretty the photo*...’

(EJB0421: I4d)

Table 5.6 presents the full inventory of adjective types included in the coding and indicates whether they were included in or excluded from the quantitative analysis.
Table 5.6. Controlled Elicitation Task: Summary of adjective coding

<table>
<thead>
<tr>
<th>Type of adjective</th>
<th>Examples</th>
<th>Quantitative analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Categorical with ser</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intensional</td>
<td>Modal/epistemic: <em>posible</em> ‘possible’, <em>necesario</em> ‘necessary’, <em>falso</em> ‘false’ Degree/quantifying: <em>completo</em> ‘complete’, <em>puro</em> ‘pure’, <em>único</em> ‘only’</td>
<td>Excluded (N=15)</td>
</tr>
<tr>
<td>Dispositional</td>
<td><em>cruel</em> ‘cruel’, <em>tonto</em> ‘stupid’, <em>cortés</em> ‘polite’, <em>discreet</em> ‘discreet’</td>
<td>Excluded (N=1)</td>
</tr>
<tr>
<td>Relational</td>
<td><em>político</em> ‘political’, <em>eléctrico</em> ‘electric’, <em>vegetariano</em> ‘vegetarian’</td>
<td>Excluded (N=19)</td>
</tr>
<tr>
<td>Kind-level</td>
<td><em>común</em> ‘common’, <em>raro</em> ‘rare’</td>
<td>Excluded (N=10)</td>
</tr>
<tr>
<td>Adverbial</td>
<td>Frequency: <em>habitual</em> ‘usual’, <em>usual</em> ‘normal’, <em>típico</em> ‘typical’, <em>constante</em> ‘constant’ Same/different: <em>diferente</em> ‘different’, <em>similar</em> ‘similar’, <em>parecido</em> ‘similar’</td>
<td>Excluded (N=60)</td>
</tr>
<tr>
<td><strong>Variable between ser and estar</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locative</td>
<td><em>cerca</em> ‘near’, <em>lejos</em> ‘far’</td>
<td>Excluded (N=9)</td>
</tr>
<tr>
<td>Polysemous</td>
<td><em>listo</em> ‘clever’ (ser)/’ready’(estar), <em>vivo</em> ‘sharp’ (ser)/’alive’ (estar)</td>
<td>Excluded (N=9)</td>
</tr>
<tr>
<td>Participle</td>
<td><em>acostumbrado</em> ‘accustomed’, <em>cocido</em> ‘cooked’, <em>reservado</em> ‘reserved’</td>
<td>Excluded (N=149)</td>
</tr>
<tr>
<td>Perfective</td>
<td><em>lleno</em> ‘full’, <em>suelto</em> ‘loose’, <em>limpio</em> ‘clean’, <em>seco</em> ‘dry’</td>
<td>Excluded (N=87)</td>
</tr>
<tr>
<td>Qualificational</td>
<td><em>bueno</em> ‘good’, <em>fuerte</em> ‘strong’, <em>picante</em> ‘spicy’, <em>suave</em> ‘soft’</td>
<td>Included (N=1124)</td>
</tr>
</tbody>
</table>

5.2.2 Quantitative analysis

5.2.2.1 Overall distribution of ser and estar
Table 5.7 shows the overall distribution of *ser* and *estar* in the data. The overall rate of *estar* selection in the data was approximately 8% versus an approximate overall rate of 92% in the selection of *ser*.

Table 5.7. Controlled Elicitation Task: Overall distribution of *ser* and *estar* with qualificational adjectives

<table>
<thead>
<tr>
<th></th>
<th><em>Ser</em></th>
<th><em>Estar</em></th>
<th><strong>Total</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1038</td>
<td>86</td>
<td>1124</td>
</tr>
<tr>
<td>%</td>
<td>92.3</td>
<td>7.7</td>
<td></td>
</tr>
</tbody>
</table>

5.2.2.2 Factor-by-factor analysis: Evidence

Next, I present the rates of selection of *ser* and *estar* according to each of the factors tested. I start with the factor evidence type, which examined the distribution of the variants across direct evidence and recalled experience contexts. I asked whether *estar* was selected at a higher frequency in contexts where direct, sensory evidence was present than in contexts of recalled experience. I predicted that speakers would show a difference in patterns of the selection of *ser* and *estar* in their responses depending on whether or not direct evidence was present in the context. Specifically, I expected participants to select *estar* more frequently in responses to contexts of direct evidence than in responses to contexts of recalled experience.

Table 5.8 shows that *estar* was used in direct evidence contexts at a rate of approximately 10% as compared with a rate of 5.5% in recalled experience contexts. On the whole, speakers selected *estar* about twice as frequently in contexts of direct evidence as compared with contexts of recalled experience, which confirmed my prediction\(^\text{16}\).

\(^{16}\) The results of a multivariate analysis of the data are reported in section 5.2.2.6, including statistical significance, relative strength and constraint ranking of factors.
Table 5.8. Controlled Elicitation Task: Distribution of *ser* and *estar* with qualificational adjectives according to evidence type

<table>
<thead>
<tr>
<th></th>
<th>Ser</th>
<th>Estar</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Direct</strong></td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>506</td>
<td>55</td>
<td>561</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>90.2</td>
<td>9.8</td>
<td></td>
</tr>
<tr>
<td><strong>Recalled</strong></td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>532</td>
<td>31</td>
<td>563</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>94.5</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1038</td>
<td>86</td>
<td>1124</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>92.3</td>
<td>7.7</td>
<td></td>
</tr>
</tbody>
</table>

5.2.2.3 Factor-by-factor analysis: Type of sensory stimuli

The second factor analysis provided a finer-grained evaluation of each of the five types of sensory stimuli (visual/auditory/olfactory/tactile/gustatory). This analysis considered whether speakers showed a difference in rates of selection of *estar* according to the sense that was active in perceiving direct evidence or recalling previous experience. The global analysis above revealed a difference in frequencies of selection of *estar* in contexts of direct evidence and recalled experience, and I anticipated that a closer look at the different types of sensory stimuli might unearth distinctions in rates of selection of *estar*, depending on the sense involved. The predictions for this variable were informed by cross-linguistic observations about the types of sensory evidence encoded by evidentials (Aikhenvald, 2004; Speas, 2004; Willett, 1988). If a language expresses evidential marking for direct sensory evidence, the visual and/or auditory senses are most likely to show a contrast with the other senses as the types of sensory input that
are distinguished by evidential markers. Hence, I expected that any contrasts in speakers’ selection of *estar* would be more pronounced in the categories of visual and/or auditory evidence.

Table 5.9 represents the distribution of *ser* and *estar* according to type of sensory stimuli. For all types of sensory evidence, *estar* appeared more frequently in direct evidence as compared with recalled experience contexts. The greatest contrast in speakers’ selection of *estar* in contexts of direct evidence versus recalled experience existed in the category of visual evidence, in keeping with the prediction. However, it must be noted that rates of the use of *estar* varied according to type of sensory stimuli, and in some categories speakers produced few to no tokens with *estar* (Direct evidence: olfactory: N=2, Recalled experience: auditory: N=1, olfactory: N=0).

Table 5.9. Controlled Elicitation Task: Distribution of *ser* and *estar* with qualificational adjectives according to evidence and type of sensory stimulus

<table>
<thead>
<tr>
<th>Type of Sensory Evidence</th>
<th>Direct</th>
<th>Recalled</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Ser</em></td>
<td><em>Estar</em></td>
<td><em>Ser</em></td>
</tr>
<tr>
<td>Visual</td>
<td>N 103</td>
<td>19</td>
<td>301</td>
</tr>
<tr>
<td>%</td>
<td>84.4</td>
<td>15.6</td>
<td>95</td>
</tr>
<tr>
<td>Auditory</td>
<td>N 45</td>
<td>5</td>
<td>46</td>
</tr>
<tr>
<td>%</td>
<td>90</td>
<td>10</td>
<td>97.9</td>
</tr>
<tr>
<td>Olfactory</td>
<td>N 74</td>
<td>2</td>
<td>48</td>
</tr>
<tr>
<td>%</td>
<td>97.4</td>
<td>2.6</td>
<td>100</td>
</tr>
<tr>
<td>Tactile</td>
<td>N 160</td>
<td>18</td>
<td>65</td>
</tr>
<tr>
<td>%</td>
<td>89.9</td>
<td>10.1</td>
<td>92.9</td>
</tr>
<tr>
<td>Gustatory</td>
<td>N 123</td>
<td>11</td>
<td>72</td>
</tr>
<tr>
<td>%</td>
<td>91.8</td>
<td>8.2</td>
<td>88.9</td>
</tr>
</tbody>
</table>
5.2.2.4 Factor-by-factor analysis: Age

In the next analysis, I evaluated the impact of the extralinguistic variable of age in the selection of *estar* in contexts of direct sensory evidence versus recalled experience. Previous studies have found that younger and less educated speakers lead the change in the semantic extension of *estar* (Cortés-Torres, 2004; Gutiérrez, 1992; Silva-Corvalán, 1986). I predicted that a similar pattern would emerge in this data. Younger (age 18-30) and less educated (high school or less) speakers were expected to show higher frequencies of selection of *estar* in contexts of direct sensory evidence than in contexts of recalled experience, as compared with older (age 50 and older) and more educated (post-secondary educated) speakers.

Table 5.10 shows that younger and older speakers produced similar patterns of selection of *ser* and *estar* in their responses to direct evidence and recalled experience contexts. Across both groups of speakers, the rates of the use of *ser* were comparable in both direct evidence and recalled experience contexts, with a ratio close to 1:1. With *estar*, however, both groups showed higher rates in direct evidence contexts. Younger speakers selected *estar* at a ratio of 2:1 in direct evidence as compared with recalled experience contexts, and this ratio was slightly lower in older speakers, at 3:2.
Table 5.10. Controlled Elicitation Task: Distribution of *ser* and *estar* with qualificational adjectives according to age and evidence type

<table>
<thead>
<tr>
<th></th>
<th>Direct</th>
<th>Recalled</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Ser</em></td>
<td><em>Estar</em></td>
<td><em>Ser</em></td>
</tr>
<tr>
<td>Younger</td>
<td>N 300</td>
<td>37</td>
<td>307</td>
</tr>
<tr>
<td></td>
<td>% 89</td>
<td>11</td>
<td>94.2</td>
</tr>
<tr>
<td>Older</td>
<td>N 206</td>
<td>18</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td>% 92</td>
<td>8</td>
<td>94.9</td>
</tr>
<tr>
<td>Total</td>
<td>N 506</td>
<td>55</td>
<td>532</td>
</tr>
</tbody>
</table>

5.2.2.5 Factor-by-factor analysis: Level of education

The results for the social variable of level of education, displayed in Table 5.11, revealed a similar pattern to that observed in the results for the variable of age. Again, rates of *ser* use were similar at an approximately 1:1 ratio in direct evidence and recalled experience contexts across both groups, whereas rates of *estar* use showed a tendency to be lower in responses to recalled experience as compared with responses to direct evidence. Less educated speakers (high school or less) selected *estar* in direct evidence over recalled experience contexts at a ratio of 3:2, and more educated speakers (with post-secondary education) at a ratio of 2:1.
Table 5.11. Controlled Elicitation Task: Distribution of *ser* and *estar* with qualificational adjectives according to level of education and evidence type

<table>
<thead>
<tr>
<th></th>
<th>Direct</th>
<th>Recalled</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Ser</em></td>
<td><em>Estar</em></td>
<td><em>Ser</em></td>
</tr>
<tr>
<td>Less educated</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>221</td>
<td>20</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>91.7</td>
<td>8.3</td>
<td>94.5</td>
</tr>
<tr>
<td>More educated</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>285</td>
<td>35</td>
<td>292</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>89.1</td>
<td>10.9</td>
<td>94.5</td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>506</td>
<td>55</td>
<td>532</td>
</tr>
</tbody>
</table>

Up to this point, the results have produced contrasting patterns of selection of *ser* and *estar* across evidence types with respect to the two social variables of age and level of education: the younger speakers and the more educated speakers used *estar* at higher rates in direct evidence contexts as compared with older speakers and less educated speakers. However, the data was not particularly robust in either case.

In sum, several general patterns emerge when we consider the factor-by-factor analysis of the data for the elicitation task. Contrasts in the frequency of selection of *ser* and *estar* were stronger across linguistic variables than social variables. For the linguistic variable of evidence type, rates of *estar* were greater in direct evidence as compared with recalled experience contexts. With respect to the different types of sensory stimuli, the data for the visual stimuli was the most robust, whereas the categories for auditory and olfactory stimuli had few or no tokens with *estar*. The social variables of age and education did not suggest strong contrasts in the use of *estar* across groups.
5.2.2.6 Multivariate analysis

The multivariate analysis included the four fixed-effect predictors Evidence, Sensory Type, Age and Education. I adjusted the levels of the predictor Sensory Type, due to the low numbers of tokens in several of the categories of sensory stimuli, as noted in the factor-by-factor analysis (section 5.2.2.3). In the regression analysis, the Visual category of sensory stimuli was maintained as a level of the predictor Sensory Type, but the categories Auditory, Olfactory, Tactile and Gustatory were combined to form one level with the label Other Sensory. I tested for all possible interactions between fixed effects in the same model (Evidence: Sensory Type, Evidence: Age, Evidence: Education, Sensory Type: Age, Sensory Type: Education, and Age: Education). Two random-effect predictors, Speaker and Adjective, were included. The results reported in Table 5.12 were taken from the best stepping-down iteration of the logistic regression\(^\text{17}\). Please see section 5.1.5 for an explanation of the terminology and format of the logistic regression analysis tables in this chapter.

\(^{17}\) The values reported for the non-significant fixed effects were taken from the full model (the first model in the step-down analysis).
Table 5.12. Mixed-effects logistic regression analysis of the contribution of external and internal factors to the probability of *estar* selection with qualificational adjectives in Buenos Aires Spanish with Speaker and Adjective as random effects

The only predictor of *estar* selection that reached statistical significance in the model was Evidence (*p* = 0.015), and none of the interactions were significant (*p*-values for all interactions were > 0.05). Contexts where direct evidence was present favoured *estar* at 0.597. Consequently, we can reject the null hypothesis that the presence of direct evidence in the context had no effect on whether speakers chose *ser* or *estar*. The predictors Sensory Type, Age and Education were
not statistically significant. The prediction that visual stimuli would be distinguished from other types of sensory stimuli in marking evidence was not statistically supported. However, it is worth noting that the constraint ranking for Sensory Type mirrored the prediction that visual evidence would be more likely than other types of sensory evidence to carry evidential marking. As for the social variables tested, the predictions with respect to age and level of education were not supported. Younger and less-educated speakers did not choose estar at significantly higher frequencies than older and more-educated speakers. If we examine the ranking of the constraints for Age and Education, the results for Age matched the expectation that younger speakers would show greater frequencies of estar selection than older speakers, but the results for Education ran contrary to expectation, with less-educated speakers producing estar at lower rates than more-educated speakers.

Based on these results, we can conclude that speakers use estar at significantly greater frequencies when direct evidence is present in the context than in contexts of recalled experience. Significantly higher rates of estar selection in direct evidence contexts occurred across speakers of different ages and levels of education, regardless of the type of sensory stimuli. This pattern suggests that the evidential use of estar with qualificational adjectives in Buenos Aires Spanish can be characterized as one of stable variation, rather than as a change in progress. In the next section, I present some examples from the data and discuss what the patterns observed therein can contribute to our understanding of the evidential use of estar.

5.2.3 Discussion: Controlled Elicitation Task

5.2.3.1 Perfective adjectives

First, I consider the variability that appears with ser and estar and perfective adjectives (e.g. seco ‘dry’, limpio ‘clean’, sucio ‘dirty’, vacío ‘empty’). In (134), the speaker alternates between ser and estar with vacía in her description of the subject la hamaca.
(134) Visual stimuli: Direct evidence stimulus: Image 7: Beach with hammock and palm trees at sunset

...muy bonita. La, la hamaca está vacía, ¿no? Porque no se, porque hay, ¿es

vacia la hamaca? No hay nadie.

‘…very beautiful. The, the hammock is estar empty, no? Because no, because there

is HABER, is SER the hammock empty? There is HABER no one there.’

(MMG0327)

Is there a difference in meaning between the sentence with estar (la hamaca está vacía) and the one with ser (¿es vacía la hamaca?)? Two possible explanations present themselves. On one hand, this may be an evidential use of estar, in which case the evidentially neutral ser clause contrasts with the estar clause that marks the speaker’s response to the presence of direct evidence. On the other hand, we can assume no meaning distinction exists, and interpret this as an instance of neutralization of the ser/estar contrast.

Contexts of ser/estar variation with perfective adjectives present counter-evidence to an approach that looks to the gradability properties of adjectives to explain the behaviour of ser and estar. The adjective vacío, in the classification of its gradability properties, is a closed-scale, absolute adjective. This class of adjectives is predicted to appear with estar in an account that views the ser/estar distinction as the spellout of the quantization of the state expressed by the adjectival predicate (see for example Gumiel-Molina & Pérez-Jiménez, 2012).

The part-structure of the subject plays a role in the quantization of the state with ser/estar predicates, according to the gradability explanation. Homogeneous subjects are predicted to appear with ser and quantized subjects with estar. However, this cannot account for the variation with the adjective sucio ‘dirty’ in (135) and (136), where the subjects la arena ‘the sand’ and el mar ‘the sea’ are both homogeneous in their part-structure.

(135) Visual stimuli: Recalled experience prompt 7: Sunsets/beaches in Argentina
...y no son muy bellas que digamos porque en sí el mar acá de Buenos Aires es muy, como está contaminado por parte del Río de la Plata, el agua no es transparente ni en pedo, es medio m-azul negrusca, y que la arena está muy sucia...

‘...and they are not very beautiful so to speak because in itself the ocean here in Buenos Aires very, since it polluted by Rio de la Plata, the water not transparent, no way, it half blackish blue, and the sand very dirty…’

(SS0328)

(136) Visual stimuli: Recalled experience prompt 7: Sunsets/beaches in Argentina

...bueno, nosotros tenemos una, un, una playa muy, o la arena es, es gruesa, marrón, el mar es marrón, es, sí, sucio...

‘...well, we have a, a, a beach very, or the sand coarse, brown, the sea brown, it is, yes, dirty…’

(CM0419)

We also observe variation with lleno ‘full/crowded’ in the examples in (137) and (138), where the two locative subjects (an unidentified city in the visual stimulus and the city of Punta del Este) encode the same part-structure and are expected to pattern consistently in verb selection, but do not.

(137) Visual stimuli: Recalled experience prompt 7: Sunsets/beaches in Argentina

...sí, Punta del Este está llena de argentinos...

‘...yes, Punta del Este full/crowded with Argentines…’
(VV0409)

(138) Visual stimuli: Direct evidence prompt 4: A dense urban landscape

...es muy descriptiva de nuestra ciudad, pero en si... es horrible, es muy llena...

‘...it is SER very descriptive of our city, but in itself... it is SER horrible, it is SER very full/crowded...’

(SS0328)

The speaker in (138) chose ser in responding to the direct evidence stimulus and the speaker in (137) used estar in responding to the recalled experience prompt, which is the opposite of the predicted pattern for the evidential use of estar. We can thus rule out the evidential use of estar as a possible explanation for these examples, and consider two alternatives. First, these data could provide support for the argument that a neutralization of the ser/estar distinction exists with perfective adjectives, and second, there could be other elements in the contexts that influence verb selection that we have not accounted for.

Nonetheless, we know that the evidential use of estar is optional, and there is evidence that it occurs in some contexts with perfective adjectives, as in (134) above. In a similar example, the speaker EJB showed variability with the adjectives calmó ‘calm’ and tranquilo ‘calm’ in combination with the subjects el mar ‘the ocean’ and el agua ‘the water’ (139-140).

(139) Visual stimuli: Direct evidence stimulus: Image 7: Beach with hammock and palm trees at sunset

...este es una calma total, el mar está calmo...

‘...this is SER a total calm, the sea is ESTAR calm...’

(EJB0421)

(140) Visual stimuli: Recalled experience prompt 7: Sunsets/beaches in Argentina

...a mí me gustan mucho las playas de Brasil en donde el agua es tranquila, como
en esta foto, es así muy muy todo muy muy, eh, uniforme el mar, y entonces vos te metés, y estás ahí, calentita...

‘...I really like the beaches in Brazil where the water is calm, like in this photo, it is like this very very all very very, uh, smooth, the ocean, and so you get in, and you are there, nice and warm…’

(EJB0421)

In the direct evidence context (139), the speaker selected estar in responding to the visual stimulus with the sentence el mar está calmo. In the recalled experience context (140), she chose ser in el agua es tranquila to describe a recalled experience of a beach in Brazil. This pattern of variability suggests an evidential use of estar.

5.2.3.2 Qualificational adjectives

Variation was present in the data with qualificational adjectives, and several patterns can be identified. First, in a pattern that coincides with the hypothesis of the evidential use of estar, there were cases where an adjective appeared with estar in responses to direct evidence stimuli and with ser in responses to recalled experience stimuli. For example, the adjective rico appeared with estar in responses to stimuli for the senses of taste and smell in the direct evidence contexts in (141) and (142) and with ser in the recalled experience context in (143).

(141) Gustatory stimuli: Direct evidence stimulus: Food item 2: Grapes

...ah, están re-ricas estas. Dulces, suaves también, la piel es delgadita, fácil de masticar.

‘...ah, these are really delicious. Sweet, mild too, the skin is thin, easy to chew.’

(CM0419)
(142) Olfactory stimuli: Direct evidence stimulus: Scent 2: Menthol/eucalyptus blend

...está riquísimo...

‘...it isestar lovely...’

(GG0411)

(143) Olfactory stimuli: Recalled experience prompt 2: Medicine (e.g. cough syrup)

...no me planteo cuando tengo que tomar un remedio si es rico, si es feo, o no. No forma parte de mi plant-, no es un planteo intelectual. Tengo que hacer esto y se acaba la milonga.

‘... I don’t consider when I have to take a remedy if it isser delicious, if it isser bad-tasting, or not. It doesn’t enter into my consider-, it isser not an intellectual consideration.

I have to do this thing and the party’s over.’

(JS0329)

The data contain multiple utterances with the adjectives bueno ‘good’, dulce ‘sweet’, frío ‘cold’, lindo ‘pretty/cute’ and rico ‘delicious’ which fall into this pattern. However, these adjectives also appeared with ser in direct evidence contexts\(^{18}\) and with estar in recalled experience contexts (at

\(^{18}\) In the utterance in (i), a response to a direct evidence stimulus, the speaker described her impressions of the stimulus in generic terms.

Gustatory stimuli: Direct evidence stimulus: Food item 1: Potato chips

(i) Muy ricas son, las papas fritas.

‘Very delicious areser, potato chips.’

(MB0331)

Many of the utterances with ser in responses to prompts for direct evidence contexts have generic meanings. A. T. Pérez-Leroux (personal communication, July 2016) observes that default generic descriptions are also reported in studies of first language acquisition where a similar methodology has been used, that is, picture elicitation tasks (Brandone & Gelman, 2009; Gelman & Brandone, 2010). In designing the elicitation task, I did not anticipate that participants would give generic descriptions of the stimuli in the direct evidence contexts. This limitation of the methodology should be taken into consideration in the creation of future experimental tasks designed to elicit the production of evidential meaning.
lower frequencies, as shown in Table 5.13), which suggests that the evidential use of estar is not obligatory, but optional.

Table 5.13. Controlled Elicitation Task: Distribution of adjectives with ser and estar by evidence type for adjectives with >1 token of estar selection in direct evidence contexts

<table>
<thead>
<tr>
<th></th>
<th>Direct evidence contexts</th>
<th>Recalled experience contexts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ser</td>
<td>Estar</td>
</tr>
<tr>
<td><strong>bueno</strong> ‘good’</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td><strong>dulce</strong> ‘sweet’</td>
<td>24</td>
<td>3</td>
</tr>
<tr>
<td><strong>feliz</strong> ‘happy’</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td><strong>frío</strong> ‘cold’</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td><strong>lindo</strong> ‘pretty/cute’</td>
<td>56</td>
<td>5</td>
</tr>
<tr>
<td><strong>rico</strong> ‘delicious’</td>
<td>72</td>
<td>4</td>
</tr>
</tbody>
</table>

In a different pattern, the adjective *picante* ‘hot/spicy’ did not appear in any direct evidence contexts with estar, but it was produced in combination with both ser and estar in recalled experience contexts only, as shown in (144) and (145), respectively.

(144) Visual stimuli: Recalled experience prompt 5: Unusual food in Argentina or that participant has tried

...las comidas típicas mexicanas son picantes...

‘…the typical Mexican foods are_{ser} spicy…’

(MC0327)
(145) Visual stimuli: Recalled experience prompt 5: Unusual food in Argentina or that participant has tried

...es como que, bueno, vayas a comer comida mexicana y digas, "esto está muy picante"...

‘...it’s like, well, you go to eat Mexican food and you say, “this is really spicy”…’

Upon closer observation, it becomes clear that the response elicited to the recalled experience prompt in (145) was a hypothetical scenario involving direct evidence, in which the use of estar is consistent with our hypothesis.

Another example of variation between ser and estar in recalled experience (but not direct evidence) contexts appears in (146) and (147), with the adjective suave ‘soft/mild’.

(146) Tactile stimuli: Recalled experience prompt 1: Hair of a pet (e.g. a cat or a dog)

...al pelo de mi mascota, es suavecito, ¿no? Para acariciar. Acariciar, mimar


‘... my pet’s hair, it is softish, no? To stroke. To stroke, to pet – very beautiful. Soft.

Pleasant.’

(MEB0330)

(147) Tactile stimuli: Recalled experience prompt 1: Hair of a pet (e.g. a cat or a dog)

...el gato le dejás afuera y le acariciás el pelo y está, está suave, viste...

‘... you leave the cat outside and you stroke its fur and it is soft, you know…’

(BB0425)
At first glance, the selection of *estar* with the qualificational adjectives *picante* and *suave* in recalled experience but not direct evidence contexts could be interpreted as counterevidence to our hypothesis about the directionality of evidential marking with *ser* and *estar*. However, it is significant that in both of these contexts, *estar* was used to refer to a sensory experience of the speaker. In (145), the speaker described the hypothetical experience of tasting Mexican food, and in (147), the speaker described the imagined sensation of touching a cat’s fur. This suggests that *estar* may be selected not only in contexts of immediate direct evidence, but also in contexts of hypothetical direct evidence to express a sensory impression.

### 5.2.3.3 *Estar bueno*: The evaluative use of *estar*

A look at the utterances containing the [*estar + bueno*] combination suggests that *estar* is selected to express evaluative meaning rather than (or in addition to) evidential meaning in these contexts. To briefly recap, evaluative meaning is closely related to evidential meaning (see section 2.6.4 for discussion). Evidential meaning makes reference to the source of the information expressed by the proposition, whereas evaluative meaning expresses a speaker’s judgment about the proposition. Evaluative meaning can also encompass the category of mirativity, which is the expression of surprise or the marking of a proposition as contrary to the expectations of the speaker.\(^{19}\)

The data for the expression [*estar + bueno*] indicate that its meaning is related to an evaluation of the subject on the part of the speaker. The adjective *bueno* (‘good, nice’) is a prototypical evaluative adjective, in the sense of the lexical semantic categories of adjectives identified by Dixon (1977, 2004) and which featured in several previous studies of the semantic extension of *estar* (see section 3.3.1 for discussion). To be clear, evaluative adjectives are a subtype of qualificational adjectives (e.g., *bueno/malo*, ‘good/bad’) whose primary lexical content is a positive or negative evaluation of a noun or proposition. For example, in the direct evidence

---

\(^{19}\) I am grateful to audiences at LSRL 2014 (University of Western Ontario) and HLS 2014 (Purdue University) for discussions of mirative meaning related to the evidential use of *estar.*
contexts (148) and (149), \([estar + bueno]\) conveyed the speaker’s (positive) evaluation of the subject.

(148) Visual stimuli: Direct evidence stimulus: Image 7: Beach with hammock and palm trees at sunset

...la hamaca paraguaya está buena, yo me imagino en la hamaca...

‘... the hammock is\textit{estar} nice, I imagine myself in the hammock…’

(NP0412)

(149) Gustatory stimuli: Direct evidence stimulus: Food item 3: Caramel candies

...igual, está muy bueno, el caramelo.

‘... anyway, it \textit{isestar} very good, the candy.’

(SM0426)

In the direct evidence context in (150) the participant (who is bald) used the expression to add a positive comment to his reaction to the touch stimulus (feathers) after making a series of negative comments, and perhaps to emphasize that the negative comments were in jest.

(150) Tactile stimuli: Direct evidence stimulus: Object 1: Bundle of feathers

...yo no me lo pondría, pero... T: ¿No? GS: No me queda, aparte no tengo

\textit{donde para agarrarlo tampoco, jajaja, pero está bueno}...

‘... I wouldn’t wear it myself, but…’ T (Researcher): ‘No?’ GS (Participant): ‘It doesn’t suit me, besides, I have nowhere to put it either, hahaha, but \textit{it isestar} nice...’

(GS0418)

The utterances in (148) through (150) were taken from direct evidence contexts, and could therefore potentially be eligible for an explanation in terms of an evidential use of \textit{estar};
however, the data also contained utterances in which [est**ar** + buen**o**] appeared in response to recalled experience prompts. The expression occurred in a description of a past experience in the recalled experience context in (151).

(151) Visual stimuli: Recalled experience prompt 7: Sunsets/beaches in Argentina

...de hecho, nosotros durante los veranos vamos a trabajar a la costa de Argentina a un lugar que se llama Miramar... está bueno... es muy lindo...

‘...actually, during the summer we go to work on the coast of Argentina to a place called Miramar... *it is**E**STAR nice... it is**S**ER very beautiful...’

(PM0327)

Finally, in the recalled experience context in (152), the participant employed [est**ar** + buen**o**] in talking about a brand of chocolate that he liked.

(152) Gustatory stimuli: Recalled experience prompt: Chocolate

...los Bon-o-Bon, esos están buenísimos...

‘... Bon-o-Bons (brand), those are**E**STAR very good...’

(BB0425)

Given the absence of direct evidence in the contexts for (151) and (152), these utterances are more accurately described as evaluative uses of est**ar** rather than evidential uses.

In the utterances where buen**o** was used in combination with est**ar**, the evaluative meaning was usually the most salient interpretation. Evaluative meaning is also present in est**ar** clauses with other qualificational adjectives, as we have seen. Based on these observations, I propose an expansion of the definition of the evidential use of est**ar** to include contexts where est**ar** is used with evaluative adjectives to convey speaker-oriented meaning.
5.2.4 Summary: Controlled Elicitation Task

If no identifiable semantic distinctions between *ser* and *estar* apply, we are forced to conclude that variable data present a neutralization of the contrast between *ser* and *estar*. However, based on the findings of the quantitative analysis and the observations of the qualitative analysis, we can reject the null hypothesis. The quantitative results indicated that the evidential use of *estar* with qualificational adjectives occurred with all types of sensory stimuli in direct evidence contexts. Significantly higher rates of *estar* selection in direct evidence contexts across speakers of different ages and levels of education suggest a pattern of stable variation, rather than a change in progress, in Buenos Aires Spanish. We can conclude that the selection of *estar* with qualificational adjectives in direct evidence contexts is an optional but statistically significant use of this verb whereby speakers encode immediate sensory experience. The qualitative analysis revealed that this use may extend to recalled experience contexts in which *estar* is chosen to describe a hypothetical scenario involving direct sensory evidence. Meanwhile, an analysis of the variable behaviour of perfective adjectives with *ser* and *estar* presented counter-evidence to an explanation of the distribution of *ser* and *estar* with adjectives in terms of gradability properties and the quantization of states. Finally, based on observations of evaluative uses of [*estar + bueno*], I proposed an expansion of the definition of the evidential use of *estar* to include contexts where *estar* appears with evaluative adjectives to produce evaluative or mirative meaning. In the next section, I apply this approach to an analysis of the interview data and elicitation data combined, with the objective of testing whether a broader characterization of the speaker-oriented use of *estar* is able to more accurately account for the patterns of variation of *ser* and *estar* with qualificational adjectives.

5.3 Interview Task

In the previous section, I argued for the expansion of the characterization of the evidential use of *estar* to include the closely related evaluative meaning that may result from the combination of
estar with an evaluative adjective. In this section, I refer to these meanings together as the speaker-oriented use of estar. I can identify two origins of the term speaker-oriented as I apply it to the use of estar with qualificational adjectives in Buenos Aires Spanish. The first source is in syntactic theories of the relationships between functional heads associated with speech act, evaluative, evidential and epistemic meanings (see for example Cinque, 1999; Faller, 2002; Roberts, 2010). The term also appears in the literature on subjectification and intersubjectification, a topic that bridges the fields of pragmatics, discourse analysis and grammaticalization. Traugott (2010) defines subjectification as “…the mechanism by which meanings are recruited by the speaker to encode and regulate attitudes and beliefs” (Traugott, 2010, p. 35). Examples of (inter)subjective expressions include speech act verbs, evaluative adjectives and adverbs, epistemic modals, evidentials, focus particles, discourse markers, and honorifics, among others (see Cuyckens et al., 2010; Davidse et al., 2010; Traugott, 2010). The use of estar that is the focus of this study falls squarely into this characterization, and I will henceforth refer to it as the speaker-oriented use of estar. The goal of this section is to evaluate the accuracy of this revised definition, using the interview data (combined with the elicitation data) to test the hypothesis that estar clauses with qualificational adjectives express speaker-oriented meaning.

5.3.1 Coding

Several challenges arose in analyzing the interview data. First, the data itself was sparse. The typically infrequent occurrence in spontaneous speech of non-phonological variables is a well-known problem in studies of variation and change (see for example Tagliamonte, 2006, p. 83). The use of estar with qualificational adjectives was no exception in this regard, as demonstrated by the overall distribution of ser and estar in the interview data in Table 5.14.
Table 5.14. Interview Task: Overall distribution of *ser* and *estar* with qualificational adjectives

<table>
<thead>
<tr>
<th></th>
<th><em>Ser</em></th>
<th><em>Estar</em></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>275</td>
<td>37</td>
<td>312</td>
</tr>
<tr>
<td>%</td>
<td>88.1</td>
<td>11.9</td>
<td></td>
</tr>
</tbody>
</table>

An additional challenge in analyzing interview data for the speaker-oriented use of *estar* was the researcher’s limited access to information about the discourse context. In the elicitation task, participants were presented with direct evidence, and the utterances they produced in these contexts could confidently be categorized as responses to direct evidence. In contrast, an analysis of the interview data using the same approach would involve guesswork on the part of the researcher. The classification of an utterance as a response to direct evidence would require the researcher to make assumptions about the nature of the evidence in the context and perhaps about the intentions of the speaker. An approach to coding that is based solely on the researcher’s judgments about the data runs the risk of reaching inaccurate conclusions.

Furthermore, the problem of ambiguity between temporal and speaker-oriented meaning in *estar* clauses underscored the need for a systematic approach to coding. Let us first consider some unambiguous examples of temporal and speaker-oriented uses of *estar* from the data. The utterances in (153) and (154) were clearly temporally-limited uses of *estar*.

(153) *...otra posibilidad era Tigre. En ese momento, Tigre no estaba tan caro...*

‘...another possibility was*Tigre* (neighbourhood). At that time, *Tigre* was*not that expensive*...’

(SM0426)

(154) *...de día es un barrio que está muy activo, demasiado activo a veces...*

‘...by day it is*a neighbourhood that is* very active, too active sometimes...’
The temporal adverbial expressions *en ese momento* (‘at that time’) and *de día* (‘by day’) anchored the temporal readings of these *estar* clauses. By contrast, the sentences in (155) and (156) were speaker-oriented uses of *estar*.

(155) *...porque el 'correo electrónico' son dos palabras, entonces 'correl', 'correl' está genial...*  
‘...because ‘electronic mail’ are two words, so ‘correl’, ‘correl’ is*great*...’

(156) *...si lo hacen en el medio de un tercero, de un privado, que es afin al gobierno, no me gusta pero no es grave. Si utilizan el canal oficial para eso, me parece que está pésimo.*  
‘...if they do it by means of a third party, a private one, that is affiliated with the government, I don’t like it but it is not serious. If they use the official channel for that, *it seems to me that it is* awful.’

The evaluative adjectives *genial* ‘great’ and *pésimo* ‘awful’ indicated speaker-oriented meaning, as did the expression *me parece que* (‘it seems to me’) in (156).

Let us now examine the meanings of the *estar* clauses in (157) and (158), where temporally-anchored and speaker-oriented readings overlapped.

(157) *...últimamente que la Argentina está demasiado violenta, muy violenta, hay mucha violencia, mucho tiro, mucho, mucha muerte, está, está muy feo, a mí me asombra.*  
‘...lately Argentina is too violent, very violent, there is a lot of violence, a lot
of shootings, a lot, a lot of death, it is \textit{very bad}, it shocks me.’

(AV0403)

(158) \ldots yo tengo mi hija, que se casó en junio, y acá junio es otoño para invierno, estuvo \textit{buenísimo}.

‘…I have my daughter, who got married in June, and here June is fall to winter, it was \textit{very nice}.’

(CC0419)

In (157), the expression \textit{está muy feo} (‘it is very bad’) was ambiguous between a temporal and an evaluative reading. On one hand, the temporal adverb \textit{últimamente} (‘lately’) could be interpreted as a temporal marker for the entire utterance. On the other hand, the expression \textit{a mí me asombra} (‘it shocks me’) and the evaluative adjective \textit{feo} (‘bad, ugly’) lent speaker-oriented meaning. Similarly, the temporal expressions in (158) (\textit{junio, y acá junio es otoño para invierno}, ‘June, and here June is fall to winter’) contrasted with the evaluative adjective \textit{bueno}, (‘good, nice’). The speaker-oriented meaning of the adjective was reinforced in this case by the emphatic suffix –\textit{ísimo}. The \textit{estar} clauses in (157) and (158) were thus ambiguous between temporal and evaluative interpretations.

As this discussion of the utterances in (153) to (158) has illustrated, the different readings generated by \textit{estar} clauses with qualificational adjectives can be classified, and ambiguous utterances identified, by means of an analysis of co-occurring elements in the sentence. This is the approach I adopted to coding. In order to test whether the temporal and speaker-oriented meanings of \textit{estar} clauses with qualificational adjectives were quantitatively significant, I coded for elements in the sentence that were predicted to appear with each of these meanings. In addition, due to the paucity of the data from the interview task, I combined the 1124 utterances from the elicitation data with the 312 utterances from the interview data (Table 5.15). The combined interview and elicitation data were sufficient for the results of a quantitative analysis to be statistically valid. However, because the interview and elicitation data were combined in this analysis, it should be noted that the input probability reported in the multivariate analysis
(see Table 5.20) was not a precise measure of the frequency of occurrence of \([estar + qualificational adjective]\) in the vernacular speech of this community. The elicitation task was designed to elicit this structure (although see section 5.2.3.2, footnote 16 for a discussion of whether the methodology was successful in achieving this objective)\textsuperscript{20}. Nonetheless, as there was no priming of the use of \textit{estar} in the elicitation task, and participants’ responses consisted of spontaneous utterances, the data from the two tasks were compatible.

Table 5.15. Overall distribution of \textit{ser} and \textit{estar} with qualificational adjectives in interview data and elicitation data combined

<table>
<thead>
<tr>
<th></th>
<th>\textit{Ser}</th>
<th>\textit{Estar}</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interview data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>275</td>
<td>37</td>
<td>312</td>
</tr>
<tr>
<td>%</td>
<td>88.1</td>
<td>11.9</td>
<td></td>
</tr>
<tr>
<td><strong>Elicitation data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1038</td>
<td>86</td>
<td>1124</td>
</tr>
<tr>
<td>%</td>
<td>92.3</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td><strong>Combined data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1313</td>
<td>123</td>
<td>1436</td>
</tr>
<tr>
<td>%</td>
<td>91.4</td>
<td>8.6</td>
<td></td>
</tr>
</tbody>
</table>

5.3.1.1 Temporal anchoring

\textsuperscript{20} The sociolinguistic interview script can also be tailored towards the elicitation of a specific form, which is another approach to collecting data when the item occurs infrequently in natural language.
Utterances that contained overt temporal anchoring were expected to produce a temporally-limited reading of an *estar* predicate. Temporal anchoring markers included non-present tense (past, future), progressive aspect, temporal adverbials (159) and subordinate clauses introduced by temporal adverbs (160).

(159)  *antes* ‘before’, *hoy* ‘today’, *a veces* ‘sometimes’, *siempre* ‘always’, *ya* ‘already’, *ahora* ‘now’, *de repente* ‘suddenly’, *después* ‘after’, *de noche* ‘at night’, *de pronto* ‘all of a sudden’, *todo el día* ‘all day’

(160)  …*lindísimo. Lindísimo. [el olor de una rosa] Cuando, cuando están maduritas, bien abiertas - es suave, ¿no? Es romántico.*

‘… beautiful. Beautiful. [the smell of a rose] *When, when they are*estar fully-grown, fully open – it is*SER pleasant, no? It is*SER romantic.’

(MEB0330:A1r)

When one or more of these markers was present, the utterance was labelled *temporally-anchored*, and when no overt temporal marking appeared, the label *not temporally-anchored* was assigned.

### 5.3.1.2 Speaker-oriented expressions

The speaker-oriented use of *estar* was predicted to occur in the presence of speaker-oriented expressions. Traugott and Dasher (2002) argue that speaker-oriented meaning, or subjectivity, “explicitly encodes SP/W’s [the speaker/writer’s] point of view, for example in deixis, modality, and marking of discourse strategies.” (2002, pp. 21-22). Subjective meaning tends to be characterized by specific linguistic expressions and principles (161).
(161) i. overt spatial, and temporal deixis,

   ii. explicit markers of SP/W attitude to what is said, including epistemic attitude to
       the proposition,

   iii. explicit markers of SP/W attitude to the relationship between what precedes and
        what follows, i.e. to the discourse structure,

   iv. the R-heuristic\textsuperscript{21} predominates.

   (Traugott & Dasher, 2002, p. 23)

Table 5.16 describes and presents examples of the elements that are coded as speaker-oriented expressions in the data. When one or more of these expressions occurred in an utterance, I assigned it the label *speaker-oriented*. I labeled utterances without speaker-oriented expressions as *non-speaker-oriented*.

\textsuperscript{21} The R-heuristic is inspired by Grice’s (1975) maxim of Relevance and Levinson’s (1983) Principle of Informativeness. Traugott and Dasher define it thus: “Say/write no more than you must, and mean more thereby.” (2002, p. 19).
Table 5.16. Summary of coding of speaker-oriented expressions

<table>
<thead>
<tr>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deictic markers (spatial and temporal)</td>
<td><em>esto</em> ‘this (one)’</td>
</tr>
<tr>
<td></td>
<td><em>aquí, ahí</em> ‘here, there’</td>
</tr>
<tr>
<td></td>
<td><em>así</em> ‘like this’</td>
</tr>
<tr>
<td></td>
<td><em>ahora</em> ‘now’</td>
</tr>
<tr>
<td>Perception verbs and semi-copula verbs</td>
<td><em>me suena</em> … ‘it sounds to me …’</td>
</tr>
<tr>
<td>(with dative pronouns)</td>
<td><em>se ve</em> … ‘it looks …’</td>
</tr>
<tr>
<td></td>
<td><em>se notaba</em> … ‘it showed …’</td>
</tr>
<tr>
<td></td>
<td><em>se me hace</em> … ‘it seems to me …’</td>
</tr>
<tr>
<td></td>
<td><em>me resulta</em> … ‘it seems to me …’</td>
</tr>
<tr>
<td>Nouns referring to perception</td>
<td><em>el sabor</em> ‘the taste’</td>
</tr>
<tr>
<td>Epistemic and evaluative adverbs</td>
<td><em>evidentemente</em> … ‘evidently’</td>
</tr>
<tr>
<td></td>
<td><em>obviamente</em> … ‘obviously’</td>
</tr>
<tr>
<td>Speaker-oriented discourse markers</td>
<td><em>para mí</em> … ‘for me …’</td>
</tr>
<tr>
<td>Verbs that express speaker attitude and/or</td>
<td><em>no me creo que</em> … ‘I don’t believe that …’</td>
</tr>
<tr>
<td>epistemic/evaluative meaning</td>
<td><em>no me gusta</em> … ‘I don’t like …’</td>
</tr>
<tr>
<td></td>
<td><em>prefiero</em> … ‘I prefer …’</td>
</tr>
<tr>
<td></td>
<td><em>me da pena cuando</em> … ‘I feel sad when …’</td>
</tr>
</tbody>
</table>
5.3.2 Quantitative analysis

5.3.2.1 Factor-by-factor analysis

I began the quantitative analysis with a look at the rates of selection of *ser* and *estar* according to each of the factors analyzed. For the factor Temporal Anchoring, I tested whether *estar* was selected at a higher frequency in contexts where a temporal marker was present than in contexts without overt temporal markers. I predicted that participants would select *estar* more frequently in utterances with temporal marking than in those without. The results in Table 5.17 confirmed that the proportion of *estar* clauses in contexts with temporal anchoring (18.4%) was more than twice that of contexts without (7.1%).

Table 5.17. Combined interview and elicitation data: Distribution of *ser* and *estar* with qualificational adjectives according to presence of temporal anchoring markers in the utterance

<table>
<thead>
<tr>
<th></th>
<th>Ser</th>
<th>Estar</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>155</td>
<td>35</td>
<td>190</td>
</tr>
<tr>
<td>%</td>
<td>81.6</td>
<td>18.4</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1158</td>
<td>88</td>
<td>1246</td>
</tr>
<tr>
<td>%</td>
<td>92.9</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1313</td>
<td>123</td>
<td>1436</td>
</tr>
</tbody>
</table>

Regarding the factor Speaker-Oriented, I examined whether speakers showed a difference in patterns of *ser/estar* selection in their responses depending on whether or not speaker-oriented expressions occurred in the utterance. Specifically, I expected participants to select *estar* more frequently in responses to contexts where speaker-oriented expressions were present. The results
in Table 5.18 show that *estar* appeared more than twice as frequently in utterances with speaker-oriented expressions (15%) as in those without (6.6%).

Table 5.18. Combined interview and elicitation data: Distribution of *ser* and *estar* with qualificational adjectives according to presence of speaker-oriented expressions in the utterance

<table>
<thead>
<tr>
<th></th>
<th>Ser</th>
<th>Estar</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>N</td>
<td>289</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>85</td>
<td>15</td>
</tr>
<tr>
<td>Absent</td>
<td>N</td>
<td>1024</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>93.4</td>
<td>6.6</td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td>1313</td>
<td>123</td>
</tr>
</tbody>
</table>

To evaluate the extent of the occurrence of *estar* clauses that were ambiguous between temporally-anchored and speaker-oriented readings, I considered the percentages of the data that corresponded to each of these factors separately, as well as the percentage of the data in which these factors overlapped. Table 5.19 presents these results. The highest percentage of the *estar* data (28.1%) was found in contexts where both temporal marking and speaker-oriented expressions were present, and the lowest percentage of the *estar* data (4.9%) corresponded to contexts where neither type of co-occurring element appeared. The percentages for unambiguous contexts of both types were similar: when an utterance featured temporal anchoring markers but no speaker-oriented expressions the percentage of *estar* clauses was 16.5%, compared with 13.6% for utterances that included speaker-oriented expressions but no overt temporal markers. Based on these findings, we can conclude that ambiguous contexts with *estar* were approximately twice as likely to occur in comparison to clauses with either a temporally-anchored or speaker-oriented interpretation.
Table 5.19. Combined interview and elicitation data: Distribution of *ser* and *estar* with qualificational adjectives according to presence of speaker-oriented expressions and temporal anchoring markers in the utterance

<table>
<thead>
<tr>
<th>Speaker-Oriented Expressions</th>
<th>Temporal Anchoring Markers</th>
<th>Ser</th>
<th>Estar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>Present</td>
<td>N 23</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>71.9</td>
<td>28.1</td>
</tr>
<tr>
<td>Present</td>
<td>Absent</td>
<td>N 266</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>86.4</td>
<td>13.6</td>
</tr>
<tr>
<td>Absent</td>
<td>Present</td>
<td>N 132</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>83.5</td>
<td>16.5</td>
</tr>
<tr>
<td>Absent</td>
<td>Absent</td>
<td>N 892</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>95.1</td>
<td>4.9</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>N 1313</td>
<td>123</td>
</tr>
</tbody>
</table>

In light of the results of the elicitation task for the social variables of age and level of education, I did not anticipate that the combined interview and elicitation data would show significant contrasts in the selection of *estar* between groups based on these external factors. Indeed, when comparing the factors age and level of education, the differences in *estar* usage according to speaker-oriented expressions and temporal marking in the context were negligible across groups (Figures 5.2 through 5.5). As such, the results of this analysis are consistent with the conclusions of previous sections that the patterns of use of *ser* and *estar* with qualificational adjectives in Buenos Aires Spanish are best characterized as stable variation.
Figure 5.1. Combined interview and elicitation data: Use of *ser* and *estar* with qualificational adjectives by age according to presence [+/-] of speaker-oriented expressions in the utterance

Figure 5.2. Combined interview and elicitation data: Use of *ser* and *estar* with qualificational adjectives by age according to presence [+/-] of temporal anchoring markers in the utterance
Figure 5.3. Combined interview and elicitation data: Use of *ser* and *estar* with qualificational adjectives by level of education according to presence [+/-] of speaker-oriented expressions in the utterance

Figure 5.4. Combined interview and elicitation data: Use of *ser* and *estar* with qualificational adjectives by level of education according to presence [+/-] of temporal anchoring markers in the utterance
5.3.2.2 Multivariate analysis

The multivariate analysis included the four fixed-effect predictors Temporal Anchoring, Speaker-Oriented Expression, Age and Education. I tested for all possible interactions between fixed effects in the same model (Temporal Anchoring: Speaker-Oriented Expression, Temporal Anchoring: Age, Temporal Anchoring: Education, Speaker-Oriented Expression: Age, Speaker-Oriented Expression: Education, and Age: Education). Two random-effect predictors, Speaker and Adjective, were included. The results reported in Table 5.20 were taken from the best stepping-down iteration of the logistic regression\textsuperscript{22}. I refer the reader to section 5.1.5 for an explanation of the format and terminology of the logistic regression analysis tables in this chapter.

\textsuperscript{22} The values reported for the non-significant fixed effects were taken from the full model (the first model in the step-down analysis).
Table 5.20. Mixed-effects logistic regression analysis of the contribution of external and internal factors to the probability of *estar* selection with qualificational adjectives in Buenos Aires Spanish with Speaker and Adjective as random effects

<table>
<thead>
<tr>
<th>Input</th>
<th>Total N</th>
<th>FW</th>
<th>Logodds</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.035</td>
<td>1436</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Significant Fixed Effects:**

**TEMPORAL ANCHORING**
- Present: 0.697, Logodds: 0.831, %: 18.4, N: 190
- Absent: 0.303, Logodds: -0.831, %: 7.1, N: 1246
- Range: 39.4

**SPEAKER-ORIENTED EXPRESSION**
- Present: 0.631, Logodds: 0.536, %: 15, N: 340
- Absent: 0.369, Logodds: -0.536, %: 6.6, N: 1096
- Range: 26.2

**Non-significant Fixed Effects:**

**AGE**
- Younger: [0.511], Logodds: 0.044, %: 9.3, N: 829
- Older: [0.489], Logodds: -0.044, %: 7.6, N: 607

**EDUCATION**
- Less: [0.522], Logodds: 0.088, %: 8.8, N: 636
- More: [0.478], Logodds: -0.088, %: 8.4, N: 800

**Random Effects:**
- Speaker: \(sd = 0.563\), N = 45
- Adjective: \(sd = 1.917\), N = 298

The two linguistic variables, Temporal Anchoring \((p < 0.001)\) and Speaker-Oriented Expression \((p < 0.001)\), were identified as highly significant predictors of *estar* selection in the model. None of the interactions were significant \((p\text{-values for all interactions were } > 0.05)\). Contexts where
overt temporal marking was present favoured estar at 0.697, which was an unsurprising result given the vast theoretical and empirical literature associating estar with temporal anchoring. However, the finding that estar was also favoured in contexts with speaker-oriented expressions at 0.631 was a novel result. This outcome was consistent with the prediction that estar would be selected to express speaker-oriented meaning.

As expected, the social variables Age and Education were not statistically significant predictors of estar selection. It is worth noting that the constraint rankings for both variables matched the predicted direction of the variation. The hierarchy for Age coincided with the expectation that younger speakers would show greater frequencies of estar selection than older speakers, and the ranking for Education corresponded to the prediction that less-educated speakers would produce estar at higher rates than more-educated speakers.

Based on these results, we can conclude that the presence of speaker-oriented expressions in the context has a highly significant effect on the selection of estar with qualificational adjectives. Overt temporal marking proves to be another highly significant contextual predictor of the use of estar. There is little to no indication of a change in progress in this community, given the similar rates of estar selection according to both of the linguistic variables across speakers of different ages and levels of education. In sum, the findings of the quantitative analysis signal stable patterns of use for temporally-anchored and speaker-oriented meanings of estar in Buenos Aires Spanish.

5.3.3 Summary and discussion: Interview Task

The previous task, the controlled elicitation task, was designed with the objective of evaluating the evidentiality hypothesis presented in Chapter 4. The interview task, by contrast, was less tightly structured, and this has permitted me to conduct an analysis that reflects a redefinition of my original hypothesis. Based on a qualitative analysis of the data in the previous section, I proposed a hypothesis of the speaker-oriented use of estar. The reformulated hypothesis claims that the [estar + qualificational adjective] structure is selected to encode speaker perspective,
which comprises a range of meanings including evidential, evaluative and mirative interpretations).

In this section, I tested the speaker-oriented hypothesis using the interview data combined with the elicitation data. A comparison of the quantitative results from the two analyses demonstrates that the speaker-oriented hypothesis provides a more accurate explanation of the patterns of _estar_ with qualificational adjectives than does the evidentiality hypothesis. The multivariate analysis of the elicitation data showed that the presence of direct evidence in the context favoured _estar_ selection with qualificational adjectives in the controlled elicitation task (FW = 0.597, Range = 19.4, \( p = 0.015 \)). For the combined interview and elicitation data, the quantitative analysis revealed that the presence of speaker-oriented expressions in the surrounding linguistic context favoured the appearance of _estar_ with qualificational adjectives (FW = 0.631, Range = 26.2, \( p < 0.001 \)). Although the linguistic predictors reached significance in both analyses, the differences in the factor weights, ranges and \( p \)-values suggest that of the two variables, the presence of speaker-oriented expressions is a more powerful predictor of _estar_ selection than the presence of direct evidence in the context. One advantage of the approach taken in the speaker-oriented analysis, which examined the linguistic markers accompanying the target structure in spontaneous speech, was that it yielded an objective measure of the variable without any guesswork on the part of the researcher. In this case, it also produced superior results when compared to the attempt to create a controlled experimental situation in the elicitation task, which led to unexpected responses from participants (see section 5.2.3.2 for discussion).

The results presented in this section are consistent with the results of the analyses of the CPT and the elicitation task in showing that linguistic factors, but not social factors, play a role in the variation in this community. We can thus conclude that the linguistic encoding of speaker perspective is a driving factor in variation in the use of _estar_ with qualificational adjectives in Buenos Aires Spanish. These findings in support of the speaker-oriented use of _estar_ represent a

---

\(^{23}\) The speaker-oriented hypothesis may also be able to account for irony, which is a meaning that can be associated with the [estar + adjective] structure (L. Colantoni, personal communication, August 2016). The utterance from the interview data in (i) is an example of the use of _estar_ with an evaluative adjective to express irony. See Gilmour, Gonzales, and Louie (2010) for a proposal on the relationship of irony and evidential constructions in English.

(i) \( \text{eh... bueno... estoy bárbaro hoy con la memoria...} \)

‘oh... well... I’m _estAR_ wonderful today with (my) memory...’
novel contribution to the broader project of capturing a more precise picture of the distribution of *ser* and *estar* with adjectives in Spanish.
The main contribution of this study is a revised characterization of the evidential use of *estar* as the speaker-oriented use of *estar*. These findings have consequences for formal theories of the elements involved, the relationships between different levels of the grammar, and the nature of semantic variation and change. In this section, I summarize the results of this study and highlight questions for further investigation.

The analysis of the speaker-oriented use of *estar* and the data from which it is drawn can contribute to debates about the syntax and semantics of *ser/estar* and adjectives. The empirical evidence does not support theories of a structural parallel between *ser* and *estar*, if we assume that speaker-oriented meaning engages levels of the grammar beyond discourse-pragmatics. A symmetrical approach to copula constructions cannot account for speaker-oriented meaning with *estar*. The evidential, evaluative, mirative and ironic interpretations that arise in certain contexts with *estar* suggest that it is more complex than *ser*. Asymmetrical approaches to *ser* and *estar*, such as those of Camacho (2012a, 2012b), Fábregas (2012) and Schmitt (1996, 2005), are compatible with the results and analysis reported here. More generally, the observation that *ser* is the only “true” copula (see for example Schmitt 1996, 2005, among others) is bolstered by data from studies of variation and change in the use of *ser* and *estar*.

The structures and meanings that are vulnerable to variation and change can signal the existence of gaps between different levels of the grammar, and the adjectives that appear in the speaker-oriented use of *estar* are one such example. Broadly, the adjectives that appear in speaker-oriented contexts with *estar* belong to the class of evaluative adjectives, but the precise boundaries of this category of adjectives remain unclear. Should the criteria for determining adjective class be based on lexical semantics, or are compositional factors involved? A gradability approach assumes the latter. Evaluative adjectives are gradable. They express polarity, as shown by the fact that they usually come in pairs of opposites (e.g. *bueno/malo* ‘good/bad’). They are a type of open-scale, relative adjective whose standard of comparison is contextually determined (see section 2.4.3 for discussion). Recent work on gradability in
adjectives focuses on vagueness as a characteristic of relative adjectives (Kennedy, 2007). Vagueness is defined as contextual variability in truth conditions, and in the case of relative adjectives, it is tied to the value of the standard of comparison. The sentence “the coffee in Rome is expensive,” contains an example of a vague predicate (Kennedy, 2007, p. 2). Kennedy argues that it is unclear whether the standard of comparison is determined by the conventional meaning of adjectives or as a function of extra-linguistic factors. He points out that,

…although it is clear that a property that we can descriptively call a ‘comparison class’ influences the computation of the standard of comparison by providing a domain relative to which this degree is computed, this property does not correspond to a constituent of the logical form. (Kennedy, 2007, p. 16).

The speaker-oriented use of estar with evaluative adjectives is an example of a context in which the feature of vagueness can be linked to variation and change. The connection between the property of vagueness and contexts of variation and change in language is an important area for future exploration.

In terms of semantic variation and change, I define the speaker-oriented use of estar in terms of subjectification. My proposal adds to the body of research on subjectification in two ways. First, the field of subjectification is a relatively recent area of study, which is still establishing its boundaries (Traugott, 2010). The exploitation of a semantic contrast in a dual copula system to encode speaker perspective is a form of subjectification not previously considered among the types of expressions that are candidates for this analysis. Is subjectification a feature of dual copula systems in other languages? Second, this study contributes data on a language other than English, which comprises the bulk of work in the field of subjectification. The hypothesis of the speaker-oriented use of estar opens the door for further investigation into subjectivity in related expressions in Spanish. As observed by Gutiérrez (1992), Camacho (2012b) and others, the speaker-oriented use of estar often involves other expressions of subjectivity such as diminutives, intensifiers and deictic markers. These types of expressions merit further examination in and of themselves.

---

25 See Traugott (2010) for an overview and discussion of investigation into linguistic contexts of subjectification.
As for the speaker-oriented use of *estar*, there is still more work to be done in delineating the aspects of meaning involved. Potential speaker-oriented meanings include speech act, evidential, evaluative and epistemic meanings. Based on data from all three tasks that suggest that *estar* may encode speaker experience with direct sensory evidence (Chapter 5), we can conclude that evidential meaning is present. In addition, the findings of the two production tasks support the claim that *estar* can be used to express evaluative meaning (sections 5.2. and 5.3.). However, the results of the contextualized preference task suggest that epistemic meaning in the form of inferred evidence can be ruled out as one of the types of speaker-oriented meaning associated with [*estar + adjective*] clauses (section 5.1.). I have not tested for speech act-related meanings in this study, and this is another area of speaker-oriented meaning with *estar* whose investigation remains outstanding.

An additional area of study that awaits is an account of the speaker-oriented meaning of *estar* in the grammar. This question is related to a formal theory of speaker-oriented meaning more generally. An account of the variation with *estar* must be compatible with a theory of speaker-oriented meaning that is cross-linguistically valid. Is the speaker-oriented use of *estar* best represented in the syntax, semantics, pragmatics, or some combination thereof? If these meanings are generated at the level of syntax, then I expect speaker-related elements may appear at one or both levels of the IP and the CP, following Blaine and Dechaine’s (2007) account of evidential morphosyntax. A theory in which speaker-oriented meaning is constructed via the syntax must explicitly address the syntax-semantics mapping. If speaker-oriented meaning is encoded at the level of semantic composition, then an account of multidimensional meaning is required to account for ambiguous sentences, possibly along the lines of Potts’s (2005) proposal of a lambda calculus for CI’s (conventional implicatures). Finally, if these meanings are located at the level of pragmatics, then models of contextually-determined meaning are needed to explain the speaker-oriented use of *estar*.

In order to distinguish which level of the grammar is involved in the representation of the speaker-oriented use of *estar*, and subjective meaning more generally, diagnostic tests represent a viable methodology. In the study of evidentials, semantic diagnostics are used to determine whether a given evidential marker operates at the illocutionary or propositional level. The various tests are compiled in Matthewson et al. (2007). From that list, Schenner (2010) identifies
five diagnostics that distinguish between illocutionary and propositional analyses of evidentials (161).

(161) a. Is the sentence felicitous if the embedded proposition is known to be false?
b. Is the sentence felicitous if the embedded proposition is known to be true?
c. Does the evidential allow speech-act readings in interrogatives?
d. Can the contribution of the evidential be questioned, doubted, rejected or disagreed with?
e. Can the evidential be embedded in constructions that do not allow illocutionary operators (the antecedent of a conditional, under a factive attitude verb, under a verb of saying)?

(2010, pp. 195-196)

Distinct patterns can be identified in the results of these diagnostic tests, which may be used to predict the semantic analysis that is appropriate for a given evidential marker. However, there are some limitations to these diagnostics. They are largely based on speaker judgments, which may be inconsistent within individual speakers and controversial across groups of speakers. Also, some of the judgments are based on subtle distinctions, which can be difficult to tease apart. Consider the contrast between an evidential marker that encodes weak speaker commitment to a strong proposition versus one that expresses strong speaker commitment to a weak proposition. Regardless, various studies that have implemented some or all of these tests find that the results of these diagnostics can point the way to an adequate semantic analysis of a given evidential. I suggest that a multi-faceted methodology, using a combination of corpus analysis, experimental tasks and judgment-based diagnostics would be a fruitful approach in future studies of this topic.
Having argued in favour of the speaker-oriented use of *estar* in Buenos Aires Spanish, it is natural to ask whether the variation in the use of *ser/estar* with adjectives in other varieties of Spanish can be accounted for in terms of this analysis. If speaker-oriented meaning is a driver of the extension of *estar* in this dialect, how can it be distinguished from other causes of variation and change, such as the loss of selectional restrictions in Los Angeles Spanish (Silva-Corvalán, 1986) or contact with English in Puerto Rican Spanish (Ortiz-López, 2000)? It would also be interesting to explore the use of *ser* and *estar* with adjectives in other Romance languages with the dual copula system. For example, Buenos Aires Spanish and the Spanish dialects of regions of Argentina north of Buenos Aires have some contact with Brazilian Portuguese. What are the consequences of the contact between Brazilian Portuguese and varieties of Argentine Spanish for the [*ser/estar* + adjective] context? Based on a history of similar paths of development of Spanish and Portuguese, as well as proposals of the syntax and semantics of [copula + adjective] contexts in Portuguese (see for example Querido, 1976; Schmitt, 1996, 2005), it is possible that an investigation into the speaker-oriented meaning of *estar* with evaluative adjectives would find evidence in support of the existence of this meaning in Brazilian Portuguese. Regardless of the language that is the focus of study, further investigations of variation and change in the [copula + adjective] context should be developed using consistent methodologies that are rooted in theories of the syntax and semantics of the elements involved.

Finally, based on the findings of this study, what does the future hold for the development of the [*ser/estar* + adjective] context in Buenos Aires Spanish? My conclusion that this speech community presents a situation of stable variation in the use of *ser* and *estar* with adjectives does not suggest we can expect any great expansion of *estar* in the near future. However, it is certain that the speaker-oriented use of *estar* has established a firm footing in this variety. I can find no better proof of this than the promotional slogan that appeared in the advertising for PRO, the political party that formed the government of Buenos Aires during the time of my field work there in 2012: *Va a estar bueno Buenos Aires.*


Camacho, J. (2012b, October). What do Spanish copulas have in common with Tibetan evidentials? Paper presented at Ser & Estar at the Interfaces, Universidad de Alcalá, Spain.


Delbecque, N. (1997). The Spanish copulas *ser* and *estar*. In M. Verspoor, K. D. Lee, & E. Sweetser (Eds.), *Lexical and syntactical construction and the construction of meaning* (pp. 247-270). Amsterdam: John Benjamins.


196


Appendix A
Stimuli: Contextualized Preference Task

1. Adriana y Juan Pablo comparten un apartamento. Son amigos desde la secundaria, y ahora son estudiantes en la universidad. Es viernes por la tarde y están conversando en la sala.
   Adriana: ¿Qué vas a hacer esta noche?
   Juan Pablo:
   A. No creo que voy a estudiar esta noche.
   B. No creo que vaya a estudiar esta noche.

   Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
   ___ La primera    ___La segunda    ___No puedo decidir

2. Una vez al mes, Adriana va a las reuniones largas y aburridas de la asociación de estudiantes en la universidad. Juan Pablo le pregunta sobre la reunión de hoy.
   Juan Pablo: ¿Qué tal la reunión hoy?
   Adriana:
   A. La reunión duraba hasta las 3:00 pero terminó a las 3:30.
   B. La reunión duró hasta las 3:00 pero terminó a las 3:30.

   Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
   ___ La primera    ___La segunda    ___No puedo decidir

3. Juan Pablo dice que tiene una noticia sorprendente. Le cuenta a Adriana que vio a su ex-novia, Julia, en la universidad, y que ella y Daniel, un chico que estaba detrás de ella durante años, iban de la mano.
   Juan Pablo: ¡No puedo creerlo!
   Adriana:
   A. ¿En serio? Yo me acuerdo que ella dijo que nunca saldría con Daniel.
   B. ¿En serio? Yo me acuerdo de que ella dijo que nunca saldría con Daniel.

   Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
   ___ La primera    ___La segunda    ___No puedo decidir

4. Juan Pablo dice que Julia y Daniel parecían muy felices juntos, pero todavía se siente un poco triste al ver a su ex-novia con otro chico.
   Adriana: ¿Te importaba mucho?
   Juan Pablo:
   A. No, solo salimos unos pocos meses.
   B. No, solo saliamos unos pocos meses.

   Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
   ___ La primera    ___La segunda    ___No puedo decidir
5. Adriana quiere levantarle el ánimo a Juan Pablo, así que le pregunta qué planes tiene para la fiesta de cumpleaños de su mejor amigo, Antonio. Juan Pablo ya tiene organizada la comida y la música, pero necesita un consejo sobre la decoración.

Adriana: ¿Por qué no conseguís unos globos?
Juan Pablo:
A. Buena idea. Todos opinan que los globos están lindos.
B. Buena idea. Todos opinan que los globos son lindos.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera    ___La segunda    ___No puedo decidir

6. Adriana le pide a Juan Pablo que le sugiera un regalo para Antonio. Juan Pablo le dice el nombre del grupo favorito de reggaetón de Antonio.

Juan Pablo: ¿Por qué no le comprás su nuevo disco a Antonio?
Adriana:
1. No, no me gustan las letras. Son malas.
2. No, no me gustan las letras. Están malas.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera    ___La segunda    ___No puedo decidir

7. Adriana le pregunta a Juan Pablo quién está invitado a la fiesta. Juan Pablo le dice que Antonio va a traer a su novia nueva.

Adriana: ¿Sabés algo de ella?
Juan Pablo:
A: Dos cosas: Antonio dijo que se conocieron bailando, y que está alta.
B: Dos cosas: Antonio dijo que se conocieron bailando, y que es alta.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera    ___La segunda    ___No puedo decidir

8. Hablan de las otras personas invitadas a la fiesta. A Adriana le alegra escuchar que invitaron a Santiago. Juan Pablo sabe que a Adriana le interesaría conocerlo mejor.

Juan Pablo: ¿Cómo van las cosas entre vos y Santiago?
Adriana:
A. Creo que le gusto, pero no estoy segura porque él es muy tímido.
B. Creo que le gusto, pero no estoy segura porque él está muy tímido.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera    ___La segunda    ___No puedo decidir

9. Adriana le pregunta quién más va a venir a la fiesta. Juan Pablo le dice que invitó a su hermanita Lucía y a sus amigos.

Adriana: ¿Y tus padres le van a dejar venir a la fiesta?
Juan Pablo:
A: Sí, pero les tuve que prometer que no **van a haber** problemas con ella y sus amigos.
B: Sí, pero les tuve que prometer que no **va a haber** problemas con ella y sus amigos.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:

___ La primera   ___ La segunda   ___ No puedo decidir

10. Juan Pablo le cuenta a Adriana que Lucía está tratando de decidir adónde ir a la escuela secundaria. Tiene que elegir entre una escuela bilingüe, **Colegio Bilingüe**, y otra escuela, **Sagrada Familia**, que ofrece una especialización en bellas artes.
Adriana: ¿Cuál prefiere Lucía?
Juan Pablo:
A. No puede decidir. Se sabe que ambas escuelas **están** muy buenas.
B. No puede decidir. Se sabe que ambas escuelas **son** muy buenas.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:

___ La primera   ___ La segunda   ___ No puedo decidir

11. Adriana pregunta dónde quedan las dos escuelas. Juan Pablo dice que **Colegio Bilingüe** queda en un área exclusiva y de clase alta, y **Sagrada Familia** queda en una zona industrial.
Juan Pablo: Si no puede decidir, puede elegir el barrio que le guste.
Adriana:
A. Entonces, el barrio en el que está **Sagrada Familia es** malo.
B. Entonces, el barrio en el que está **Sagrada Familia está** malo.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:

___ La primera   ___ La segunda   ___ No puedo decidir

12. Adriana piensa que Lucía debería ir al **Colegio Bilingüe** por su ubicación. Le pregunta a Juan Pablo qué opina.
Juan Pablo: Creo que va a seguir a sus amigos.
Adriana:
A: Sí, es importante que ella **tiene** amigos en la escuela.
B: Sí, es importante que ella **tenga** amigos en la escuela.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:

___ La primera   ___ La segunda   ___ No puedo decidir

13. Adriana le dice a Juan Pablo que tiene una noticia emocionante. Le cuenta que acaba de comprar el pasaje a Barcelona para el mes próximo.
Juan Pablo: ¡Buenísimo! ¿Dónde te vas a quedar allí?
Adriana:
A. Me voy a quedar en lo de mi tía; su casa **está** grande.
B. Me voy a quedar en lo de mi tía; su casa **es** grande.
14. Juan Pablo dice que no sabía que Adriana tenía una tía en Barcelona. Adriana le dice que se llama Beti, es la hermana menor de su madre, y que todo el mundo dice que Beti y Adriana se parecen.

Juan Pablo: ¿Cómo se parecen?
Adriana:
A: Ella es alta, como yo.
B: Ella está alta, como yo.

15. Juan Pablo le pregunta a Adriana si ella y su tía se parecen en su carácter también.
Adriana: Sí, los de nuestra familia dicen que somos parecidas en nuestro carácter.
Juan Pablo:
A: Está tímida, entonces (se ríen).
B: Es tímida, entonces (se ríen).

16. Juan Pablo le pregunta a Adriana si tiene planes de visitar Madrid durante su viaje. Adriana dice que le preocupa el clima de Madrid en julio.
Adriana: Como todo el mundo sabe, hace mucho calor en Madrid en julio.
Juan Pablo:
A: Sí, y es muy seco también.
B: Sí, y está muy seco también.

17. Juan Pablo quiere saber qué planes tiene Adriana para su viaje a Barcelona. Adriana le dice que quiere hacer un recorrido por la ciudad, visitar los lugares de interés, e ir de compras. De repente, Adriana parece preocupada.
Juan Pablo: ¿Qué te pasa?
Adriana:
A: Me preocupa que mi valija chica no me alcance.
B: Me preocupa de que mi valija chica no me alcance.

18. Juan Pablo tiene una valija más grande y ofrece prestársela a Adriana. Va a buscarla y se la muestra a Adriana.
Juan Pablo: Mirá, te puedo prestar esta.
Adriana:
   A: Gracias! **E**sta **l**inda.
   B: Gracias! **E**s **l**inda.

*Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:*
   ___ La primera    ___ La segunda    ___ No puedo decidir

19. Juan Pablo pregunta qué tipo de compras Adriana piensa hacer en Barcelona. Adriana le dice que quiere comprar zapatos. Le gusta el diseño de los zapatos españoles.
Juan Pablo: Espero que lleves mucha plata con vos!
Adriana:
   A. No, no tanto. Según Beti, los zapatos **son** baratos allí.
   B. No, no tanto. Según Beti, los zapatos **están** baratos allí.

*Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:*
   ___ La primera    ___ La segunda    ___ No puedo decidir

20. Juan Pablo le dice a Adriana que quiere pedirle un favor. Quiere saber si ella puede comprarle una cámara en España y traerla, si él le da el dinero.
   Adriana: Por supuesto!
   Juan Pablo:
      A. En general, las cámaras no **son** baratas aquí.
      B. En general, las cámaras no **están** baratas aquí.

*Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:*
   ___ La primera    ___ La segunda    ___ No puedo decidir

21. Adriana le pregunta a Juan Pablo si sabe la marca y el modelo de cámara que quiere. Él le dice que estaba leyendo en internet sobre una cámara nueva que tiene todas las características que él quiere.
Adriana: Espero que quepa en la valija!
   Juan Pablo:
      A. Creo que la cámara **es** chica porque cabe en el bolsillo.
      B. Creo que la cámara **está** chica porque cabe en el bolsillo.

*Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:*
   ___ La primera    ___ La segunda    ___ No puedo decidir

22. Ni Juan Pablo ni Adriana tienen ganas de cocinar esta noche, así que deciden salir a cenar. Juan Pablo le pregunta a Adriana si hay un restaurante donde quiere ir.
   Juan Pablo: ¿Sabés dónde querés ir?
   Adriana:
      A. A algún lugar en el centro, se dice que todos los restaurantes del barrio **están** horribles.
B. A algún lugar en el centro, se dice que todos los restaurantes del barrio son horribles.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera ___La segunda ___No puedo decidir

23. Van a una zona en el centro con muchos restaurantes y dan vueltas buscando un lugar para cenar. Juan Pablo le pregunta a Adriana si conoce algún restaurante bueno en la zona.
Adriana: No me acuerdo, hace mucho que no vengo a esta zona.
Juan Pablo:
A. La última vez que vine a este barrio, habían muchas fiestas porque era Carnaval.
B. La última vez que vine a este barrio, había muchas fiestas porque era Carnaval.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera ___La segunda ___No puedo decidir

24. Deciden que quieren sentarse afuera. Se encuentran en una calle donde algunos restaurantes tienen patios. Adriana se para al lado de la puerta de un restaurante que se llama Las Flores.
Adriana: ¿Qué te parece este lugar?
Juan Pablo:
A. Bien, el patio es grande.
B. Bien, el patio está grande.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera ___La segunda ___No puedo decidir

25. Les gusta el patio, pero Adriana no está segura si le va a gustar la comida. Adriana lee el menú mientras Juan Pablo mira a la gente cenando en el patio.
Adriana: ¿Qué pensás?
Juan Pablo:
A. Creo que la comida está rica acá.
B. Creo que la comida es rica acá.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera ___La segunda ___No puedo decidir

Adriana: ¿Pedimos el vino Termidor?
Juan Pablo:
A. No, como todo el mundo sabe, es horrible.
B. No, como todo el mundo sabe, está horrible.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:

203
27. Hablan de qué van a pedir. Juan Pablo tiene mucha hambre pero Adriana comió mucho en el almuerzo, así que decide pedir una ensalada.

Juan Pablo: ¿Querés compartir un aperitivo?
Adriana:
A. Sí. ¿Pedimos queso? **Está** rico.
B. Sí. ¿Pedimos queso? **Es** rico.

*Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:*

___ La primera    ___La segunda    ___No puedo decidir

28. Quedan en que van a compartir el queso como aperitivo. Adriana le pregunta a Juan Pablo qué va a pedir.

Juan Pablo: Voy a pedir el lomo. ¿Vas a pedir una ensalada, es todo?
Adriana:
A. ¿Sabés? **Cambié** de opinión. Voy a pedir el risotto.
B. ¿Sabés? **He cambiado** de opinión. Voy a pedir el risotto.

*Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:*

___ La primera    ___La segunda    ___No puedo decidir

29. Finalmente el mozo llega a tomar el pedido, pero no pide disculpas por el retraso, ni les sonríe. Anota el pedido rápidamente y se apura a alejarse de la mesa.

Adriana: ¿Qué te parece el mozo?
Juan Pablo:
A. No **es** muy amable.
B. No **está** muy amable.

*Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:*

___ La primera    ___La segunda    ___No puedo decidir

30. Adriana le pregunta a Juan Pablo si trabajó en un restaurante alguna vez. Juan Pablo dice que no, pero le dice a Adriana que su tío es dueño de un restaurante.

Adriana: ¿Se llevan bien, vos y tu tío?
Juan Pablo:
A. Sí, él **está** muy amable.
B. Sí, él **es** muy amable.

*Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:*

___ La primera    ___La segunda    ___No puedo decidir

31. Terminan el aperitivo y luego llegan los platos principales. A Adriana le gusta el risotto, pero Juan Pablo parece decepcionado con el lomo.

Adriana: ¿Qué te pasa?
Juan Pablo:
A. El lomo es seco.
B. El lomo está seco.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera  ___La segunda  ___No puedo decidir

32. Desde donde están sentados en el patio, pueden oír la música en vivo de algún lugar cercano. Juan Pablo le pregunta a Adriana si sabe el nombre del instrumento que se toca, y ella le dice que suena como una mandolina.

Juan Pablo: ¿Qué es una mandolina?
Adriana:
A. Se parece a una guitarra, pero está chica.
B. Se parece a una guitarra, pero es chica.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera  ___La segunda  ___No puedo decidir

33. Terminan la cena y hablan de qué van a hacer después. Juan Pablo sugiere que vayan al cine.

Adriana: Buena idea! ¿Querés ver Los Vengadores?
Juan Pablo:
A. No, ya la he visto.
B. No, ya la vi.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera  ___La segunda  ___No puedo decidir

34. Adriana le sugiere que vean El Caballero Oscuro. Juan Pablo le cuenta que intentó verla dos veces y se agotaron las entradas, pero sus amigos Gonzalo y Rodrigo consiguieron entradas.

Adriana: ¿Qué les pareció la película?
Juan Pablo:
A. Dijeron que es muy buena.
B. Dijeron que está muy buena.

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera  ___La segunda  ___No puedo decidir

35. Deciden que van a intentar ver El Caballero Oscuro, y empiezan a caminar hacia el cine. Llegan 20 minutos tarde. Adriana pide dos entradas en la taquilla.

Juan Pablo: ¿Qué dijo?
Adriana:
A. Es tarde, ya no se venden entradas. ¿Querés volver a casa?
B. Es tarde, ya no se venden entradas. ¿Querés volver a casa?

Si las dos son aceptables, y si Ud. fuera a decir esto, ¿cómo lo diría?:
___ La primera  ___La segunda  ___No puedo decidir

205
Appendix B
Language and Personal Background Questionnaire

Información personal

- Nombre: ______________________   Apellido: ______________________
- Sexo:       Mujer       Hombre
- Año de nacimiento:_____________________________________________________
- Lugar de nacimiento:________________________________________________________________________
- Ocupación:________________________________________________________________________________
- ¿En qué vecindario vive?_____________________________________________________________________
- ¿Hace cuánto tiempo vive allí?_________________________________________________________________
- ¿En qué otros barrios vivió? Indique el nombre del barrio y los años que vivió allí.
- Nivel de escolaridad:
  Primaria   Secundaria   Vocacional   Universitaria
Indique el nombre de la escuela o del instituto educativo y los años que asistió.
Información lingüística

- ¿Cuál es su lengua materna?

- ¿Cuál es la lengua materna de:
  - su madre?
  - su padre?

- ¿Aprendió su lengua materna desde su nacimiento?

Si no, explique por favor:

- ¿Habla otros idiomas? ¿Cuál(es)?

Indique el idioma y el nivel de su capacidad en cada idioma.

<table>
<thead>
<tr>
<th>Idioma</th>
<th>¿A qué edad lo aprendió?</th>
<th>¿Dónde lo aprendió?</th>
<th>Nivel de capacidad</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>En casa/Escuela</td>
<td>Básico/Intermedio/Avanzado/Superior</td>
</tr>
<tr>
<td></td>
<td></td>
<td>En casa/Escuela</td>
<td>Básico/Intermedio/Avanzado/Superior</td>
</tr>
<tr>
<td></td>
<td></td>
<td>En casa/Escuela</td>
<td>Básico/Intermedio/Avanzado/Superior</td>
</tr>
<tr>
<td></td>
<td></td>
<td>En casa/Escuela</td>
<td>Básico/Intermedio/Avanzado/Superior</td>
</tr>
</tbody>
</table>

- ¿Hay algo más que quiere compartir sobre su historia lingüística y personal que puede ser relevante a este estudio?
Table C.1. Participant Biographical Data

<table>
<thead>
<tr>
<th>Code</th>
<th>Sex</th>
<th>Age</th>
<th>Place of birth</th>
<th>Occupation</th>
<th>Neighbourhood of residence</th>
<th>Level of education (highest)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB0425</td>
<td>M</td>
<td>18</td>
<td>Buenos Aires</td>
<td>Student</td>
<td>San Fernando</td>
<td>Secondary</td>
</tr>
<tr>
<td>EG0403</td>
<td>F</td>
<td>21</td>
<td>Buenos Aires</td>
<td>Domestic Worker</td>
<td>Villa Devoto</td>
<td>Secondary</td>
</tr>
<tr>
<td>VC0420</td>
<td>F</td>
<td>23</td>
<td>Grand Bore</td>
<td>Homemaker</td>
<td>Barrio Alve</td>
<td>Secondary complete</td>
</tr>
<tr>
<td>SS0328</td>
<td>F</td>
<td>24</td>
<td>Móron Buenos Aires</td>
<td>Electrician/ Singer</td>
<td>La Boca</td>
<td>Secondary</td>
</tr>
<tr>
<td>JB0423</td>
<td>M</td>
<td>24</td>
<td>Buenos Aires</td>
<td>Sales</td>
<td>Caballito</td>
<td>Secondary</td>
</tr>
<tr>
<td>ALB0420</td>
<td>F</td>
<td>25</td>
<td>Grand Bore Bs As</td>
<td>Homemaker</td>
<td>Santa Monica</td>
<td>Primary</td>
</tr>
<tr>
<td>FZ0322</td>
<td>F</td>
<td>27</td>
<td>Buenos Aires</td>
<td>Singer/ Service worker</td>
<td>Villa Crespo</td>
<td>Secondary</td>
</tr>
<tr>
<td>CR0420</td>
<td>F</td>
<td>27</td>
<td>Martenidad de Santa Rosa</td>
<td>Homemaker</td>
<td>Los Frenos</td>
<td>Secondary</td>
</tr>
<tr>
<td>MP0322</td>
<td>M</td>
<td>29</td>
<td>Argentina</td>
<td>Musician/ Service worker</td>
<td>Villa Crespo</td>
<td>Secondary</td>
</tr>
<tr>
<td>MCC0417</td>
<td>F</td>
<td>29</td>
<td>San Nicolas</td>
<td>Administrative Assistant</td>
<td>Montserrat</td>
<td>Secondary</td>
</tr>
<tr>
<td>MT0407</td>
<td>M</td>
<td>30</td>
<td>Buenos Aires</td>
<td>Factory Worker</td>
<td>Belgrano</td>
<td>Secondary</td>
</tr>
<tr>
<td>CP0331</td>
<td>F</td>
<td>20</td>
<td>Buenos Aires</td>
<td>Student</td>
<td>Villa Crespo</td>
<td>University</td>
</tr>
<tr>
<td>VV0409</td>
<td>F</td>
<td>20</td>
<td>Buenos Aires</td>
<td>Student</td>
<td>Ciudadela</td>
<td>University</td>
</tr>
<tr>
<td>ARG0410</td>
<td>F</td>
<td>21</td>
<td>Buenos Aires</td>
<td>Student</td>
<td>Villa Crespo</td>
<td>University</td>
</tr>
<tr>
<td>NR0417</td>
<td>F</td>
<td>21</td>
<td>Buenos Aires</td>
<td>Student</td>
<td>Remedios de Escalada, Lanús</td>
<td>University</td>
</tr>
<tr>
<td>EP0331</td>
<td>M</td>
<td>25</td>
<td>Buenos Aires</td>
<td>Student</td>
<td>Villa Crespo</td>
<td>University</td>
</tr>
<tr>
<td>NP0412</td>
<td>F</td>
<td>26</td>
<td>Buenos Aires</td>
<td>Photographer/ Designer</td>
<td>San Martin</td>
<td>Vocational</td>
</tr>
<tr>
<td>SF0321</td>
<td>F</td>
<td>27</td>
<td>Buenos Aires</td>
<td>Student</td>
<td>Palermo</td>
<td>University &amp; Vocational</td>
</tr>
<tr>
<td>ES0411</td>
<td>M</td>
<td>27</td>
<td>Córdoba</td>
<td>Systems Engineer</td>
<td>Boedo</td>
<td>University</td>
</tr>
<tr>
<td>AM0412</td>
<td>F</td>
<td>27</td>
<td>Buenos Aires</td>
<td>Student</td>
<td>San Telmo</td>
<td>University</td>
</tr>
<tr>
<td>CM0419</td>
<td>F</td>
<td>27</td>
<td>Buenos Aires</td>
<td>Teacher</td>
<td>Floresta</td>
<td>University</td>
</tr>
<tr>
<td>MEB0330</td>
<td>F</td>
<td>28</td>
<td>Capital Federal</td>
<td>Broadcaster</td>
<td>Colegiales</td>
<td>Vocational</td>
</tr>
<tr>
<td>GG0411</td>
<td>F</td>
<td>28</td>
<td>Buenos Aires</td>
<td>Broadcaster</td>
<td>Floresta</td>
<td>Vocational</td>
</tr>
<tr>
<td>FZ0320</td>
<td>M</td>
<td>30</td>
<td>Buenos Aires</td>
<td>Systems Analyst/ Information Technologist</td>
<td>Almagro</td>
<td>University</td>
</tr>
</tbody>
</table>

208
<table>
<thead>
<tr>
<th>Code</th>
<th>Sex</th>
<th>Age</th>
<th>Place of birth</th>
<th>Occupation</th>
<th>Neighbourhood of residence</th>
<th>Level of education (highest)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF0421</td>
<td>F</td>
<td>30</td>
<td>Buenos Aires</td>
<td>University Instructor</td>
<td>Villa Crespo</td>
<td>University &amp; Vocational</td>
</tr>
<tr>
<td>CC0403</td>
<td>F</td>
<td>50</td>
<td>Buenos Aires</td>
<td>Domestic Worker</td>
<td>Villa Devoto</td>
<td>Primary</td>
</tr>
<tr>
<td>SM0426</td>
<td>F</td>
<td>50</td>
<td>Buenos Aires</td>
<td>Homemaker</td>
<td>San Fernando</td>
<td>Secondary</td>
</tr>
<tr>
<td>NM0405</td>
<td>M</td>
<td>51</td>
<td>Buenos Aires</td>
<td>Theatre Set Designer</td>
<td>José Mármol</td>
<td>Secondary incomplete</td>
</tr>
<tr>
<td>MSG0410</td>
<td>F</td>
<td>51</td>
<td>Lomas de Zamora</td>
<td>Preschool Teacher</td>
<td>Villa Crespo</td>
<td>Secondary</td>
</tr>
<tr>
<td>JS0329</td>
<td>M</td>
<td>55</td>
<td>Argentina</td>
<td>Artist</td>
<td>Villa Crespo</td>
<td>Primary</td>
</tr>
<tr>
<td>AV0403</td>
<td>M</td>
<td>58</td>
<td>Avellaneda, Buenos Aires</td>
<td>Orderly</td>
<td>Liniers</td>
<td>Secondary</td>
</tr>
<tr>
<td>GS0418</td>
<td>M</td>
<td>58</td>
<td>Buenos Aires</td>
<td>Administrative Assistant</td>
<td>Castelar Provincia de Bs As</td>
<td>Secondary</td>
</tr>
<tr>
<td>AP0426</td>
<td>M</td>
<td>59</td>
<td>Buenos Aires</td>
<td>Retirée/ Superintendent</td>
<td>San Fernando</td>
<td>Primary</td>
</tr>
<tr>
<td>CC0419</td>
<td>F</td>
<td>63</td>
<td>Buenos Aires</td>
<td>Dressmaker</td>
<td>Monte Castro</td>
<td>Secondary</td>
</tr>
<tr>
<td>MB0331</td>
<td>F</td>
<td>96</td>
<td>Buenos Aires</td>
<td>Homemaker</td>
<td>Caballito</td>
<td>Primary</td>
</tr>
<tr>
<td>PM0327</td>
<td>F</td>
<td>50</td>
<td>Arg. Capital Federal</td>
<td>Singer/ Actress</td>
<td>Villa Crespo</td>
<td>Vocational</td>
</tr>
<tr>
<td>CN0405</td>
<td>F</td>
<td>52</td>
<td>Argentina</td>
<td>Business Administrator</td>
<td>Colegiales</td>
<td>University</td>
</tr>
<tr>
<td>LG0421</td>
<td>F</td>
<td>53</td>
<td>Florida, Provincia de Bs As</td>
<td>Business Administrator</td>
<td>San Martín</td>
<td>Vocational</td>
</tr>
<tr>
<td>LL0423</td>
<td>F</td>
<td>54</td>
<td>Montevideo, Uruguay</td>
<td>Sales &amp; Marketing</td>
<td>Almagro</td>
<td>University &amp; Vocational</td>
</tr>
<tr>
<td>MMG0327</td>
<td>F</td>
<td>55</td>
<td>Lomas de Zamora</td>
<td>Actress/ Broadcaster</td>
<td>José Mármol</td>
<td>Vocational</td>
</tr>
<tr>
<td>JR0417</td>
<td>M</td>
<td>55</td>
<td>Buenos Aires</td>
<td>Accountant</td>
<td>Villa Crespo</td>
<td>University</td>
</tr>
<tr>
<td>EJB0421</td>
<td>F</td>
<td>55</td>
<td>Buenos Aires</td>
<td>High School Teacher</td>
<td>Almagro</td>
<td>University</td>
</tr>
<tr>
<td>LP0404</td>
<td>F</td>
<td>58</td>
<td>Buenos Aires</td>
<td>Administrative Assistant</td>
<td>Flores</td>
<td>Vocational</td>
</tr>
<tr>
<td>MC0327</td>
<td>F</td>
<td>77</td>
<td>Tandil (Pcia. de Buenos Aires)</td>
<td>Retirée/ Journalist</td>
<td>Recoleta</td>
<td>Vocational</td>
</tr>
<tr>
<td>EP0404</td>
<td>M</td>
<td>87</td>
<td>Buenos Aires</td>
<td>Retirée/ History Professor</td>
<td>Montserrat</td>
<td>University</td>
</tr>
<tr>
<td>Code</td>
<td>First language</td>
<td>Mother's first language(s)</td>
<td>Father's first language(s)</td>
<td>Other language(s) (proficiency level)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>---------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB0425</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EG0403</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate), Italian (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VC0420</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS0328</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>German (intermediate), English (intermediate), Portuguese (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JB0423</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALB0420</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FZ0322</td>
<td>Spanish</td>
<td>Greek &amp; Spanish</td>
<td>Spanish</td>
<td>English (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CR0420</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MP0322</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Italian</td>
<td>English (intermediate), Italian (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCC0417</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MT0407</td>
<td>Spanish</td>
<td>Spanish &amp; Italian</td>
<td>Spanish</td>
<td>English (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP0331</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VV0409</td>
<td>Spanish</td>
<td>Spanish &amp; Guarani</td>
<td>Spanish</td>
<td>French (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARG0410</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate), Portuguese (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR0417</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EP0331</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP0412</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF0321</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English; age 8; school; advanced</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ES0411</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate), French (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AM0412</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (basic), Portuguese (basic), Italian (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CM0419</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate), Portuguese (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEB0330</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG0411</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (basic), French (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FZ0320</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF0421</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Portuguese (intermediate), English (intermediate), Italian (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC0403</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SM0426</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM0405</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSG0410</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Belgian</td>
<td>Spanish</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JS0329</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Italian</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AV0403</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS0418</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP0426</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC0419</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>First language</td>
<td>Mother's first language(s)</td>
<td>Father's first language(s)</td>
<td>Other language(s) (proficiency level)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>----------------</td>
<td>----------------------------</td>
<td>----------------------------</td>
<td>---------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MB0331</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PM0327</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN0405</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate), Italian (basic), Portuguese (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LG0421</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LL0423</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>French (intermediate), Italian (basic), Portuguese (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMG0327</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JR0417</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EJB0421</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (basic), Italian (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LP0404</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (basic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC0327</td>
<td>Spanish</td>
<td>Spanish</td>
<td>Spanish</td>
<td>English (intermediate)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EP0404</td>
<td>Spanish</td>
<td>Spanish &amp; Basque</td>
<td>Spanish</td>
<td>English (basic), French (basic)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix D
Stimuli: Controlled Elicitation Task

Figure D1. Practice Image

Image Courtesy of freedigitalphotos.net: “Young Asian Businessman”: kongsky
Figure D2. Image 1

Image Courtesy of freedigitalphotos.net: “The Balkans”: Evgeni Dinev

Figure D3. Image 2

Image Courtesy of freedigitalphotos.net: “Cute Dog”: photostock
Figure D4. Image 3

Image Courtesy of freedigitalphotos.net: “Girl kidding as a rabbit”: koratmember

Figure D5. Image 4

Image Courtesy of freedigitalphotos.net: “New York Storm”: Damian Brandon
Figure D6. Image 5

Image Courtesy of freedigitalphotos.net: “Chapulines con guacamole”: angelo

Figure D7. Image 6

ImageCourtesy of freedigitalphotos.net: “Colorful taxis”: think4photop
Figure D8. Image 7

Image Courtesy of freedigitalphotos.net: “Relaxing Beach”: winnond
Figure D9. Image 8

Image Courtesy of freedigitalphotos.net: “Strawberry Cake”: piyato
Figure D10. Object 1

Figure D11. Object 2

Figure D12. Object 3
Appendix E
Script: Guided Interview Task

1. El barrio

Quisiera saber algo de su barrio. ¿Cómo se llama? ¿Desde cuándo/hace cuánto tiempo que vive allí? ¿Me lo podría describir un poco?

¿Por qué se mudó (su familia) a ese barrio? ¿Por trabajo?
¿Vive(n) lejos de donde trabaja(n) (Ud./sus padres)?
¿Vive(n) cerca del resto de su familia y amigos?

¿Cuánto hace que su familia vive en Buenos Aires? ¿De dónde es su familia?

¿Se acuerda de cuándo llegó a su barrio por primera vez? Quisiera que me contara algo sobre cuándo llegó y cómo le fue.

¿Ha notado algún cambio en el barrio en el tiempo que ha vivido allí? ¿Cómo ha cambiado el barrio?

¿Su barrio es un lugar donde los vecinos se conocen? ¿Conoce Ud. a sus vecinos? ¿Me podría contar algo de ellos?

Se dice que en estos tiempos los vecinos no se conocen, no se hablan tanto como en el pasado. ¿Qué opina Ud.? ¿Por qué cree Ud. que ha cambiado esto?

¿Tiene amigos que viven cerca? ¿Dónde viven? Cuando se junta con ellos, ¿qué (clase de actividades) hacen juntos?
¿Hay un lugar en su barrio donde la gente se reúne? ¿Adónde van a tomar café o una copa? ¿Adónde van de mañana? ¿De tarde? ¿Hay un café local que frecuenta Ud.?

¿Dónde se reúne la gente fuera de la casa?

¿Qué le gusta más de su barrio? ¿Qué cosas no le gustan de su barrio?

2. El pasado

¿Ha pasado algún acontecimiento en su barrio? Por ejemplo, un incendio, un crimen, un accidente, etc. ¿Dónde? ¿Lo vio Ud.?

¿Cómo respondieron los vecinos en el momento? ¿Ayudaron a las víctimas de alguna manera con comida, ropa, techo?

¿Llegó la policía? ¿Hubo una investigación?

¿Se acuerda Ud. de (acontecimiento histórico)? ¿Dónde estuvo cuando se enteró de la noticia? ¿Qué hizo cuando supo de (acontecimiento)?

¿Cómo le afectó a Ud. y a su familia? ¿Y a sus amigos? ¿Y a su barrio?

3. La comida familiar

¿La familia comía junta todos los días? ¿Se comía algo especial una vez por semana? ¿Cuándo? ¿Qué comían?

¿Hubo un plato especial que preparaba su madre/padre/abuela/abuelo/tía/etc.?
¿Qué les gustaba comer a los miembros de su familia?

¿Le gusta cocinar? (¿Hornear?) ¿Qué le gusta cocinar (u hornear)?

¿Hay una diferencia entre la comida hoy y en el pasado? ¿Por qué cree Ud. que es así?

4. Los remedios caseros

¿Ud. va al médico cuando se enferma? ¿Qué hace Ud. cuando se enferma? ¿Toma alguna medida de prevención ante la gripe o las enfermedades?

¿Conoce Ud. algún remedio casero que se usaba en el pasado cuando alguien se enfermaba? ¿Cuáles remedios caseros se usan todavía hoy? ¿Cuáles son los remedios caseros que conoce Ud.? ¿Se cura todavía el empacho? ¿Cómo? ¿Conoce a alguien que haya tenido haciérselo curar?

5. Los años escolares

Quisiera saber algo de sus tiempos de colegio. ¿A qué escuela primaria asistió? ¿Dónde fue a la escuela secundaria?

El edificio de la escuela, ¿cómo lo describiría? ¿Qué tipo de escuela era? ¿Cuántos alumnos había en la escuela o en su clase?

¿Cuáles materias estudió? ¿Tenía Ud. una materia preferida, o una que no le gustaba?

¿Llevaban uniforme en esa escuela? ¿Me lo podría describir?
Si no llevaban uniforme, ¿cómo se vestían las chicas? ¿Los chicos? ¿Cómo lo compara con como se visten los chicos hoy?

¿Fue alguna vez a un paseo escolar? ¿Se fue de viaje de egresados? ¿Adónde fue? ¿Pasó algo interesante?

6. Los pasatiempos/los medios/el entretenimiento

¿Tiene Ud. algún pasatiempo? ¿Juega algún deporte? ¿Qué hace?
¿Cómo empezó a hacer eso?
¿Entró en una competencia alguna vez? ¿Qué pasó? ¿Ganó?

¿Toca Ud. un instrumento? ¿Cuál? ¿Cuánto hace que toca ____?
¿Cómo empezó a tocar ________?
Si no, ¿hay un instrumento que quisiera aprender a tocar? ¿Por qué?

¿Pasa mucho tiempo en sitios sociales de Internet como Facebook, Twitter, etc.? ¿Chatea mucho en línea? ¿Cuánto tiempo pasa en línea?

¿Le gusta jugar videojuegos? ¿Cuál es su videojuego favorito? ¿Cómo se juega? ¿Cuál fue su mejor resultado?

¿Qué tipo de música le gusta? ¿Qué música escucha estos días? ¿Ha ido a un concierto últimamente? ¿Qué le pareció?

¿Cuál es el mejor concierto que ha visto?
¿Quién es su cantante o grupo favorito? ¿Por qué?

¿Tiene Ud. una película preferida? ¿De qué se trata?
7. Los chicos y los padres hoy en día

Mucha gente dice que hay una diferencia entre los chicos hoy y los chicos en el pasado. ¿Está de acuerdo? ¿Por qué? ¿Cuál es la diferencia?

¿Podría comparar cómo se divirtieron los chicos hace # años con cómo se divierten los chicos hoy?

Mucha gente dice que hay una diferencia entre los padres hoy y los padres en el pasado. ¿Qué opina?

¿Hay cosas que hacen los miembros de su familia a veces que le molestan a Ud.?

¿Le gastó una broma a su hermano o a su hermana una vez?

¿Cuál es la cosa más cómica que ha hecho?

8. Los viajes

¿Ha viajado mucho? ¿Adónde fue? ¿Por cuánto tiempo? ¿Pasó algo interesante?

A veces la gente tiene con problemas en los aeropuertos. Por ejemplo, se pierde el equipaje, se cancela o se atrasa el vuelo, problemas de comunicación, etc. ¿Le pasó algo a Ud. alguna vez?

¿Le ha pasado algo cómico por no hablar o entender un idioma?

¿Cuál es la cosa más interesante que le ha pasado mientras viajaba?

¿Hay algún lugar que no conoce y que quisiera visitar? ¿Por qué? ¿Qué sabe de este lugar?
9. Las costumbres


¿Había ciertas costumbres o tradiciones en estas fiestas cuándo las festejaba en familia? ¿Qué hacían? ¿Dónde las festejaban? ¿Qué comían?

¿Tiene un recuerdo de (una Navidad) en particular?

¿Había otras costumbres en su familia que se acuerda de su niñez? ¿Otros días festivos que pasaban en familia?

10. Misceláneo

¿Dónde más ha vivido? ¿Me podría contar algo de ese lugar? ¿Por cuánto tiempo vivió allí?

¿Ha conocido a alguien famoso una vez? ¿Quién? ¿Dónde? ¿Habló con la persona?

¿Ha hecho algún trabajo voluntario? ¿Me podría contar un poco de eso?

¿Tiene o ha tenido una mascota? ¿Me la podría describir? ¿Puede o podía hacer algunos trucos?

11. El idioma

¿Ha notado algo interesante de cómo habla la gente de acá?

Se dice que el castellano ha cambiado y sigue cambiando mucho. ¿Qué opina?
¿Ha notado algún cambio en la manera de hablar de la gente de acá?

¿Cómo se sabe que alguien es de acá cuando habla? ¿La gente de distintos barrios habla distinto?

¿La gente joven y la gente mayor, habla diferente? ¿Cuál? ¿Nota alguna diferencia en la manera de hablar entre Ud. y sus (padres/hijos)?

¿Ud. habla igual que sus amigos? ¿Cuáles diferencias se notan?

¿Qué opina de cómo habla la gente joven de hoy? ¿Qué ha cambiado y qué no ha cambiado?