Variation and change in Spanish future temporal expression: Rates, constraints, and grammaticalization

Jessica Elana Aaron

Follow this and additional works at: https://digitalrepository.unm.edu/span_etds

Part of the European Languages and Societies Commons, and the Latin American Languages and Societies Commons
THE UNIVERSITY OF NEW MEXICO
ALBUQUERQUE, NEW MEXICO 87131

POLICY ON USE OF THESES AND DISSERTATIONS

Unpublished theses and dissertations accepted for master’s and doctor’s degrees and deposited in the University of New Mexico Library are open to the public for inspection and reference work. They are to be used only with due regard to the rights of the authors. The work of other authors should always be given full credit. Avoid quoting in amounts, over and beyond scholarly needs, such as might impair or destroy the property rights and financial benefits of another author.

To afford reasonable safeguards to authors, and consistent with the above principles, anyone quoting from theses and dissertations must observe the following conditions:

1. Direct quotations during the first two years after completion may be made only with the written permission of the author.

2. After a lapse of two years, theses and dissertations may be quoted without specific prior permission in works of original scholarship provided appropriate credit is given in the case of each quotation.

3. Quotations that are complete units in themselves (e.g., complete chapters or sections) in whatever form they may be reproduced and quotations of whatever length presented as primary material for their own sake (as in anthologies or books of reading) ALWAYS require consent of the authors.

4. The quoting author is responsible for determining “fair use” of material he uses.

This thesis/dissertation **Jessica Elana Aaron** has been used by the following persons whose signatures attest their acceptance of the above conditions. (A library which borrows this thesis/dissertation for use by its patrons is expected to secure the signature of each user.)

<table>
<thead>
<tr>
<th>NAME AND ADDRESS</th>
<th>DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Jessica Elana Aaron

Spanish and Portuguese

This dissertation is approved, and it is acceptable in quality and form for publication on microfilm:

Approved by the Dissertation Committee:

[Signatures]

Chairperson

Accepted: Amy B. Wohlert

Dean, Graduate School

NOV 28 2005

Date
VARIATION AND CHANGE IN SPANISH FUTURE TEMPORAL EXPRESSION: RATES, CONSTRAINTS, AND GRAMMATICALIZATION

BY

JESSICA ELANA AARON

B.A., Spanish/Political Science, Stanford University, 1999
M.A., Latin American Studies, Stanford University, 2000
M.A., Anthropology, University of New Mexico, 2002

Dissertation

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
Spanish and Portuguese

The University of New Mexico
Albuquerque, New Mexico

December 2006
DEDICATION

To my elders with disabilities who fought for the passage of the 1974 Individuals with Disabilities Education Act (IDEA) and the 1990 Americans with Disabilities Act (ADA), without which I would not have had access to this world.
ACKNOWLEDGMENTS

I have had the fortune of working with an exceptionally talented, supportive and dedicated committee to whom I offer my deepest and most sincere gratitude. A Conception Company Company, gracias por dedicar tu tiempo, tus ideas lúcidas y tus palabras calurosas a esta alumna casi desconocida, y por haberme ayudado a mantener la perspectiva internacional en mi trabajo. To Melissa Axelrod, thank you for always being there and for helping me to believe in my abilities. To Catherine Travis, for keeping interaction and real speakers in the picture, and for the fun times. To Joan Bybee, thank you for sharing your brilliance and warmth, and for always expecting the best from me. A Rena Torres Cacoullos, por haberme iluminado el mundo de lingüística y mucho más, por tu exigencia y tu amistad, por todo, gracias, que no hay palabras.

I would like to thank the Latin American and Iberian Institute at the University of New Mexico for its financial support throughout my career at UNM, including a Title VI Fellowship, a Field Research Grant, and three years of Ph.D. Fellowship, the last of which made possible the timely completion of this dissertation. Also for their financial support, I would like to thank the Spanish Colonial Research Center and Joseph P. Sánchez for having me as a part of the team, and the Department of Linguistics at the University of New Mexico for allowing me to teach my dream course twice.

At the Department of Spanish and Portuguese at the University of New Mexico, I would like to thank Rosario Johnson, Martha Hurd and Kate Merril for administrative crisis intervention, and Professors Maria Dolores González, Kathryn Joy McKnight, Leila Lehnman, Kimberlé López, and Clancy Clements for their guidance and support. Thanks to my colleagues Matthew Alba, José Esteban Hernández and Esther Brown, who went before me and helped to shed light on the occasionally hazy path toward the Ph.D. Thanks also to Susan Buescher, Ana Medina-Murillo, Sonia Balasch, Michelle Salazar, Patricia Cano, and Jessica Cano, who helped make sure I survived when support systems and technology fell apart; and to my personal care attendants, particularly Ali, Audrey, Clara, and Darla, for helping me live my life independently. At the Spanish Colonial Research Center, Angélica Sánchez-Clark, thanks for being a great supervisor and friend, and Elsa Delgado, por compartir tu comida extraña y vino fino. At the University of Florida, thank you to my new colleagues and students for cheering me on in the final stages.

Thanks also go to my small family: to my mother Gigi, for letting me fly away so young and so far, to my little brother Blake, for looking up to me, to my grandmother Ellen, for making me laugh and visiting me often, and to my aunt Rhonda, for always asking about my "book" and listening to the answer. Thanks also to Anna Iris for keeping me young and giving me hope for the future. To my even smaller local family, my cats Fili, Puma and Clarice, thanks for keeping me warm and for filling my house with joy.
VARIATION AND CHANGE IN SPANISH FUTURE TEMPORAL EXPRESSION: RATES, CONSTRAINTS, AND GRAMMATICALIZATION

BY

JESSICA ELANA AARON

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
Spanish and Portuguese

The University of New Mexico
Albuquerque, New Mexico

December 2006
Variation and change in Spanish future temporal expression
Jessi Elana Aaron

BA, Spanish and Political Science, Stanford University, 1999
MA, Latin American Studies, Stanford University, 2000
MA, Anthropology (Ethnology), University of New Mexico, 2002
PhD, Spanish and Portuguese (Hispanic Linguistics), University of New Mexico, 2006

This quantitative, diachronic study of variation between the Spanish Synthetic Future
*cantaré* (SF) and the Periphrastic Future *voy a cantar* (PF) tackles the development of
these two expressions within and outside the realm of future temporal reference in
Spanish since Old Spanish through the early 21st century. Working within the framework
of grammaticization and variation theory, this study begins with a qualitative and
quantitative form-based analysis of each form, based on over 5,500 tokens extracted from
a 935,000-word written and oral corpus.

In the case of the SF, the relationship between SF occurrences in future temporal
contexts and those in non-future epistemic contexts is addressed, and it is shown that
these two contexts of use have statistically significant differences in distributional
tendencies. In the case of the PF, which appeared with a relative frequency of just over
10% in the 17th-century data, it is argued that very few uses are not subsumed under
future temporal expression.

The form-based analyses are followed by four independent function-based
analyses for each time period spanning the 17th through the 20th centuries, in which the
factors conditioning SF-PF variation are examined. A comparison of these analyses
reveals, first, a contextual generalization of the PF into erstwhile SF territory beginning
in the 17th century, and second, a shift in the division of labor in future temporal
expression as the SF loses and the PF gains default future status in the 20th century. A
return to the form-based study offers insight into this shift, suggesting that the increased use of SF in epistemic modal contexts (e.g. ¿dónde estará María ahora? 'Where might María be now?') perturbed the division of labor between the SF and PF, thereby altering the linguistic conditioning on the two future temporal expression variants in Spanish today.

The form-based analyses offer a diachronic portrait of each expression and the particular contexts in which each expression occurs. The fuction-based analyses give a snapshot of the division of labor between these two variants at each time period, while comparison of these analyses reveals the details of the dynamic process of the side-by-side grammaticization of these variants.
Table of Contents

List of Figures ......................................................................................................................... xi
List of Tables .......................................................................................................................... xii
1 Introduction ......................................................................................................................... 1
  1.1 The Spanish future ............................................................................................................ 1
    1.1.1 Two constructions ....................................................................................................... 1
    1.1.2 Romance futures: two questions ................................................................................ 5
      1.1.2.1 Semantic differences and speaker choice ............................................................ 6
      1.1.2.2 Epistemic modality and SF semantics .................................................................. 11
  1.2 A usage-based approach .................................................................................................. 15
    1.2.1 Grammaticization ....................................................................................................... 19
    1.2.2 Linguistic correlates of grammaticization .................................................................. 23
    1.2.3 Methods in a usage-based framework ....................................................................... 24
  1.3 Variation theory .............................................................................................................. 28
    1.3.1 Structured variability and competition ..................................................................... 28
    1.3.2 Frequency and language change .............................................................................. 30
  1.4 The present study ........................................................................................................... 32
2 Methodology ....................................................................................................................... 34
  2.1 Corpus ............................................................................................................................ 34
  2.2 Data extraction and exclusions ...................................................................................... 35
  2.3 Coding and hypotheses .................................................................................................... 39
    2.3.1. Verb class ................................................................................................................. 41
    2.3.2 Temporal modification and specificity ..................................................................... 42
    2.3.3 Grammatical person and animacy ............................................................................ 43
    2.3.4 Clause type ............................................................................................................... 46
    2.3.5 Sentence type ............................................................................................................ 46
    2.3.6 Polarity ....................................................................................................................... 47
    2.3.7. Lexical type .............................................................................................................. 48
  2.4 The variable context ........................................................................................................ 49
    2.4.1 Neutralization of meaning in discourse .................................................................... 49
    2.4.2 Defining the envelope of variation .......................................................................... 51
  2.5 Variable rule analysis ..................................................................................................... 53
3 The Synthetic Future ........................................................................................................... 56
  3.1 Origins and development ............................................................................................... 56
  3.2 Changes in form: Increased fixedness and syntactic variability ..................................... 67
    3.2.1 Increased fixedness .................................................................................................... 67
    3.2.2 Syntactic variability I: cantar lo hé ............................................................................ 69
    3.2.3 Syntactic variability II: cantarélo ............................................................................ 73
  3.3 The functions of SF ........................................................................................................ 77
    3.3.1 Futurity and epistemicity ........................................................................................... 77
    3.3.2 Distinguishing meanings .......................................................................................... 81
  3.4 Characterizing the SF ..................................................................................................... 85
    3.4.1 Temporal uses of SF ............................................................................................... 85
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2</td>
<td>Epistemic SF</td>
<td>90</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Other proposed SF functions</td>
<td>97</td>
</tr>
<tr>
<td>3.5</td>
<td>Differentiating temporal and epistemic SF</td>
<td>104</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Grammatical person and subject animacy</td>
<td>104</td>
</tr>
<tr>
<td>3.5.2</td>
<td>Clause and sentence type, polarity</td>
<td>112</td>
</tr>
<tr>
<td>3.6</td>
<td>Summary</td>
<td>120</td>
</tr>
<tr>
<td>4</td>
<td>The Periphrastic Future</td>
<td>124</td>
</tr>
<tr>
<td>4.1</td>
<td>Origins</td>
<td>124</td>
</tr>
<tr>
<td>4.2</td>
<td>Changes in form: Increased fixedness</td>
<td>126</td>
</tr>
<tr>
<td>4.3</td>
<td>The functions of PF</td>
<td>129</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Temporal PF</td>
<td>129</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Other proposed PF functions</td>
<td>131</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Non-temporal PF</td>
<td>134</td>
</tr>
<tr>
<td>4.4</td>
<td>The grammaticization of the PF</td>
<td>135</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Verb class</td>
<td>138</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Lexical diffusion</td>
<td>141</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Subject</td>
<td>145</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Locatives</td>
<td>148</td>
</tr>
<tr>
<td>4.5</td>
<td>Temporal adverbials</td>
<td>153</td>
</tr>
<tr>
<td>4.6</td>
<td>Summary</td>
<td>154</td>
</tr>
<tr>
<td>5</td>
<td>Forms in competition</td>
<td>155</td>
</tr>
<tr>
<td>5.1</td>
<td>Language variation in use</td>
<td>155</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Grammaticization and language change</td>
<td>156</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Form-function asymmetry in the Spanish future</td>
<td>163</td>
</tr>
<tr>
<td>5.2</td>
<td>Results: Frequency</td>
<td>170</td>
</tr>
<tr>
<td>5.3</td>
<td>Variable rule analyses</td>
<td>172</td>
</tr>
<tr>
<td>5.3.1</td>
<td>The 17th century</td>
<td>174</td>
</tr>
<tr>
<td>5.3.2</td>
<td>The 19th century</td>
<td>179</td>
</tr>
<tr>
<td>5.3.3</td>
<td>The 20th century, written data</td>
<td>187</td>
</tr>
<tr>
<td>5.3.4</td>
<td>The 20th century, spoken data</td>
<td>192</td>
</tr>
<tr>
<td>5.3.5</td>
<td>The face of the two Futures: 300 years</td>
<td>198</td>
</tr>
<tr>
<td>5.3.5.1</td>
<td>Subject: measuring volition and intention</td>
<td>205</td>
</tr>
<tr>
<td>5.4</td>
<td>Summary and discussion</td>
<td>210</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Semantics and variant choice: a diachronic perspective</td>
<td>210</td>
</tr>
<tr>
<td>5.4.2</td>
<td>(In)stability and renewal</td>
<td>214</td>
</tr>
<tr>
<td>6</td>
<td>Synchronized change</td>
<td>218</td>
</tr>
<tr>
<td>6.1</td>
<td>Unresolved questions</td>
<td>218</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Beyond the variable context</td>
<td>218</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Semantic generalization and functional space</td>
<td>224</td>
</tr>
<tr>
<td>6.2</td>
<td>Frequency</td>
<td>228</td>
</tr>
<tr>
<td>6.3</td>
<td>Changing magnitudes of effect</td>
<td>230</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Stative verbs</td>
<td>231</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Interrogatives</td>
<td>235</td>
</tr>
<tr>
<td>6.3.3</td>
<td>Temporal adverbials</td>
<td>244</td>
</tr>
<tr>
<td>6.4</td>
<td>Overall tendencies: grammaticization, variation and divergence</td>
<td>249</td>
</tr>
<tr>
<td>6.5</td>
<td>Summary of results and implications: shifting defaults</td>
<td>253</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>7</td>
<td>Summary and conclusions</td>
<td>257</td>
</tr>
<tr>
<td>7.1</td>
<td>The approach</td>
<td>257</td>
</tr>
<tr>
<td>7.2</td>
<td>The analysis</td>
<td>258</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Form-based studies</td>
<td>258</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Function-based study</td>
<td>259</td>
</tr>
<tr>
<td>7.3</td>
<td>Explaining change</td>
<td>261</td>
</tr>
<tr>
<td>7.4</td>
<td>Recommendations for further research</td>
<td>262</td>
</tr>
<tr>
<td>7.5</td>
<td>Choosing a future</td>
<td>263</td>
</tr>
<tr>
<td>Appendix: Corpus</td>
<td></td>
<td>268</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>271</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1.1. *Force dynamic in the generation of a grammaticization* 22
Figure 2.2. *Contextual features coded* 40
Figure 3.3. *Rate of intervening pronouns with temporal SF, 13th-17th centuries* 68
Figure 3.4. *Relative frequency of epistemic SF uses by data set, percentage of all SF* 92
Figure 4.5. *Rate of intervening material with temporal PF by data set* 128
Figure 5.6. *Relative frequencies of temporal PF and epistemic SF by data set* 228
Figure 6.7. *Proportion of stative verbs in temporal PF and temporal SF* 233
Figure 6.8. *Proportion of epistemic SF in all SF, compared to proportion of epistemic SF in all occurrences of stative verbs in SF* 233
Figure 6.9. *Rates of co-occurring temporal adverbials in PF and temporal SF by data set* 246
Figure 6.10. *Direction of effect for overt temporal specification/modification, interrogatives and stative verbs on temporal SF, epistemic SF and PF* 256
List of Tables

Table 2.1. Absolute and relative frequencies of PF and SF by century, raw and normalized per 10,000 words  
Table 2.2. Exclusions from general study, token counts  
Table 2.3. Exclusions from function-based study, token counts: epistemic SF and motion/habitual PF  
Table 3.4. Relative frequency of analytic cantar lo hé among SF tokens with co-occurring pronouns, 13th and 15th centuries  
Table 3.5. Relative frequency of cantarélo among synthetic SF tokens with co-occurring pronouns, 17th century  
Table 3.6. Distribution of subject in temporal SF by data set  
Table 3.7. Distribution of verb class in temporal SF by data set  
Table 3.8. Distribution of clause type in temporal SF by data set  
Table 3.9. Distribution of sentence type in temporal SF by data set  
Table 3.10. Distribution of polarity in temporal SF by data set  
Table 3.11. Distribution of subject in epistemic SF by data set  
Table 3.12. Distribution of verb class in epistemic SF by data set  
Table 3.13. Distribution of clause type in epistemic SF by data set  
Table 3.14. Distribution of sentence types in epistemic SF by data set  
Table 3.15. Distribution of polarity in epistemic SF by data set  
Table 3.16. Proportion of concessive SF within epistemic SF by data set  
Table 3.17. Proportion of 1st-person singular and inanimate singular subjects in temporal (T) and epistemic (E) SF by data set  
Table 3.18. Distribution of kinds of inanimate subjects in temporal and epistemic SF  
Table 3.19. Distribution of verb class with singular animate (A) and singular inanimate (I) subjects in all SF by data set  
Table 3.20. Distribution of verb class with singular animate and singular inanimate subjects, all data sets  
Table 3.21. Relative frequency of temporal and epistemic SF by clause type, all data sets  
Table 3.22. Head verbs of object complement clauses with SF, 20th-century speech  
Table 3.23. Distribution of sentence type in temporal and epistemic SF by data set  
Table 3.24. Relative frequency of temporal and epistemic SF by sentence type, all data sets  
Table 3.25. Distribution of polarity in temporal and epistemic SF by data set  
Table 3.26. Relative frequency of temporal and epistemic SF by polarity, all data sets  
Table 4.27. Occurrence of non-future motion PF within all PF by data set  
Table 4.28. Distribution of verb class in temporal PF by data set  
Table 4.29. Most frequently occurring lexical types in all PF by data set  
Table 4.30. Type frequency and type-token ratio in all PF by data set  
Table 4.31. Distribution of subject in temporal PF by data set  
Table 4.32. Distribution of locatives with all PF by data set  
Table 4.33. Distribution of types of clause-internal locative in all PF by data set
Table 4.34. Association of clause-internal locatives with trajectory motion in PF by data set  
Table 4.35. Rate of occurrence of temporal adverbials with all PF by data set  
Table 5.36. Absolute and relative frequencies of temporal PF and SF by century, raw and normalized per 10,000 words  
Table 5.37. Variable rule analysis of the contribution of factors to occurrence of PF, 17th century  
Table 5.38. Variable rule analysis of the contribution of factors selected as significant to occurrence of PF, 19th century  
Table 5.39. Distribution of kinds of subordinate clauses in temporal SF and PF, 19th century  
Table 5.40. Head verbs of object complement clauses with SF and PF, 19th century  
Table 5.41. Variable rule analysis of the contribution of factors selected as significant to occurrence of PF, 20th-century written data  
Table 5.42. Distribution of interrogatives in PF and temporal SF by data set  
Table 5.43. Variable rule analysis of the contribution of factors selected as significant to occurrence of PF, 20th-century spoken data  
Table 5.44. Heads of object complement clauses with temporal SF, epistemic SF and PF, 20th-century speech  
Table 5.45. Variable rule analyses of the contribution of factors selected as significant to occurrence of PF, 17th-20th centuries  
Table 5.46. Factors' direction of effect by data set  
Table 5.47. Ordering of magnitude of effect by data set, in descending order  
Table 5.48. Distribution of verb classes in PF and temporal SF by data set  
Table 5.49. Rate of singular animate subjects in PF and temporal SF by data set  
Table 5.50. Rate of first-person singular subject in temporal SF and PF by data set  
Table 5.51. Rate of singular inanimate subjects in temporal SF and PF by data set  
Table 6.52. Distribution of verb class in PF and temporal SF by data set  
Table 6.53. Proportion of interrogatives with PF, temporal SF and epistemic SF by data set  
Table 6.54. Distribution of temporal and epistemic SF by sentence type, all data sets  
Table 6.55. Distribution of subject within interrogative clauses, 17th century  
Table 6.56. Distribution of subject within interrogative clauses, 19th century  
Table 6.57. Distribution of subject within interrogative clauses, 20th-century writing  
Table 6.58. Distribution of subject within interrogative clauses, 20th-century speech  
Table 6.59. Variable rule analyses of the contribution of factors selected as significant to occurrence of PF, 20th-century spoken data  
Table 6.60. Relative frequency of temporal and epistemic SF according to temporal specificity, all data sets  
Table 6.61. Relative frequency of temporal and epistemic SF according to stativity, all data sets  
Table 6.62. Relative frequency of temporal and epistemic SF according to sentence type, all data sets  
Table 7.63. Factors favoring the choice of SF vs. 'go'-based future* in Romance languages and English will
1. Introduction

1.1 The Spanish future

El futuro es un tiempo verbal cargado de matices significativos y valores modales que exceden en mucho de la simple referencia a un momento o periodo posterior al que se describe, o aquel en que se habla. Estos valores pueden incluso anular la futuridad misma...

- Salvador Fernández Ramírez (1986:284)

Salvador Fernández Ramírez' introductory words to his treatment of the Spanish future in his Gramática Española are not an exaggeration. They reflect the complexity of futurity in Spanish and the constructions used to express it, representing a rich, multifaceted territory of Spanish morphosyntax that has captivated linguists for over fifty years. In this study, I join these linguists, aiming to contribute to our knowledge of the scope of the Spanish future.

1.1.1 Two constructions

In Modern Spanish, there are two constructions that are most commonly associated with future meaning. The older of these two forms is the Synthetic Future (SF) (sometimes called the Morphological Future, Simple Future (e.g. Cartagena 1995-1996), Grammatical Future (e.g. Matte Bon 2005), or simply Future (e.g. Butt & Benjamin 1994)), shown in (1.1). The second of these is the Periphrastic Future (PF) (sometimes called the Analytic Future (e.g. Bersch in 1986)), in (1.2).

(1.1) Luego, sobre las nueve y media estaré en la reunión. (COREC, CPCON006A, 20s)
'Later, around nine-thirty I will be (SF) in the meeting'

(1.2) Yo voy a confessar aquí un secreto (COREC, CECON023A, 20s)
'I am going to confess (PF) a secret here'
A diachronic look at future temporal reference in Spanish reveals even more possibilities within the semantic domain. The Latin future, as in *cantabo, dicam*, appears only in the earliest texts, and was replaced by the SF (Company 2003:10). Other ways of expressing futurity since the 13th century have included other expressions with *habere* 'have', such as *he de cantar*¹ and *cantar lo hé* (Company 1985-1986; Company & Medina Urrea 1999), and the futurate Present (Cartagena 1995-1996), among others.² Example (1.3), taken from a 20th-century Peninsular speech sample, presents some of the alternatives speakers use to refer to future time without using the SF or the PF.

---

¹ Note that in the 20th-century speech data in the present study, this construction was not used for future temporal reference. In fact, this construction only appeared twice in the entire 20th-century spoken corpus, in the following obligation contexts: i) Bueno, esa es la gran frase de aquel gran juez inglés, ¿no? que dijo que para ser juez, primero se ha de ser honrado, imparcial, tener un poco de sentido común y no va mal saber un poco de derecho. 'Well, that's the great saying of that great English judge, right?, who said that to be a judge; first one should be (*haber de* + INF) honest, impartial, have a bit of common sense and it doesn't hurt to know a little law' (COREC, CCCON012A, 20s); ii) La tercera hemos de dársela a los educadísimos rutherford, que pincharon - furgonetas; 'We should give (*haber de* + INF) the third one to the oh-so-polite newspaper deliverymen, who slashed – stationwagons' (COREC, CACON034A, 20s).

² For example, the subjunctive (*quiero que cantes mañana* 'I want you to sing tomorrow') and certain periphrases with infinitives (*debo/puedo/quiero/pienso cantar mañana* 'I should/can/want to sing tomorrow') (Sedano 1994:225n1; see also Cartagena (1995-1996:81) for a similar list; and Lope Blanch 1983 [1972] for a description of options for future temporal expression in Mexico City speech). Sedano also includes the imperative (*canta mañana* 'sing tomorrow!') in this list (1994:225n1).
Uy pues yo cuando las \textit{pida} (Pres. Subj.), el mes de julio o de agosto Yo también \textit{puedo hacer} (\textit{poder} + INF) intermedios, Pues tengo pensado o bien , o Turquía, uno de los dos, no estoy todavía muy convencida. Estoy casi más convencida para - por irme a - a Turquía, pero depende de la salida del - del vuelo del avión. ¿Entiendes? Pero ¿del día o la hora? No - no, el día ya lo sé Vamos a ver, es que tengo que hacer (\textit{tener que} + INF) una trampa , yo quiero \textit{pedir} (\textit{querer} + INF) mis vacaciones del quince de julio al quince de agosto. El quince de julio que empieza a contar desde el lunes que es día quince, luego yo tengo (P) todo el viernes por la tarde, todo el sábado y todo el domingo que las agrego a mis vacaciones, más luego yo me tengo que \textit{venir} (\textit{tener que} + INF) a trabajar, se supone, el día diecisiete de agosto, que es viernes, y \textit{aparezco} (P) el lunes que es dieciocho. (COREC, CBCON007A, 20s)

'Oh, well, when I \textit{request} (Pres. Subj.) [my vacation], in the month of July or August, I \textit{can do} (\textit{poder} + INF) it in between, too. Well I have planned either Turkey, or one of the two, I'm still not really convinced. I'm almost more sure about going to, to Turkey, but it depends on the flight departure. See? But the day or the time? No – no, I already know the day. Let's see, the thing is, I have to do (\textit{tener que} + INF) a trick, I \textit{want to request} (\textit{querer} + INF) my vacation from July 15 through August 15. July 15, that starts to count from Monday which is the 15\textsuperscript{th}, then I have (P) all Friday afternoon, all Saturday and all Sunday that I add onto my vacation, but then I have to \textit{come} (\textit{tener que} + INF) to work, supposedly, on the 16\textsuperscript{th} of August, which is Friday, and I \textit{appear} (P) on Monday, which is the 18\textsuperscript{th}.'

In (1.3), the speaker is describing vacation plans, which will include cheating a bit in order to have an extra weekend of vacation. In describing the events that will take place during the plan, the speaker refers to these future events using: i) the Present Subjunctive: \textit{pida} 'request'; ii) \textit{poder} 'be able' + INF: \textit{puedo hacer} 'I can do'; iii) \textit{tener que} 'have to' + INF: \textit{tengo que hacer} 'I have to do' and \textit{tengo que venir} 'I have to come'; iv) \textit{querer} 'want' + INF: \textit{quiero pedir} 'I want to request'; and v) the Present (P): \textit{tengo} 'I have' and \textit{aparezco} 'I appear'.

As Lope Blanch notes, many of these options carry modal meaning along with future meaning (1983 [1972]); most have yet to be studied systematically as Spanish future temporal reference constructions. The use of the futurate Present, however, has been examined by Gutiérrez (1995), and the analytic SF \textit{cantar lo hêtre} versus synthetic
*cantaré* in Medieval Spanish has been the subject of some controversy. Some scholars have seen the synthetic SF *cantaré* and the analytic SF *cantar lo hē* as simply alternate ways of saying the same thing, related diachronically (Butler 1969:178; Rossi 1975:394; C. Lyons 1978:227; Fleischman 1982:69-70; Alvar y Portier 1983:247-249; Stengaard 1985:225-227; Kitova 1986:415; Schwegler 1990:125; Andrés-Suárez 1994:107-108), while some have argued that *cantar lo hē*, which has an intervening pronoun (*lo*) between the infinitive and *habere*, developed special meaning at some point in its diachronic trajectory (see Benveniste 1968; Company 1985-1986, 2006; Company & Medina Urrea 1999). Company and Medina Urrea (1999; see also Company 2006), for instance, in an examination of Medieval Spanish, found that *cantar lo hē* performed a focalizing function, which indicated the salience of the pronoun’s referent, soon before this construction’s demise in the 17th century.3

Most studies on the Spanish future have focused on the two grammatical constructions that compete in most varieties today (or at least in the past century): the SF

---

3 For the present study, I have included both constructions under the label of SF, and have treated them as a group. Subsequent quantitative analysis of the variation between the Analytic SF and the Synthetic SF in 15th-century data revealed marked differences in distribution of these two constructions, which supports Company’s (1985-1986, 2006) finding that these constructions did not have the same function. Multivariate analyses of yet another construction, *cantarélo*, which in these data reaches its peak of relative frequency in the 17th century (3%, N=38/1329), revealed statistically significant differences with *lo cantaré*, differences which parallel the differences found between *cantar lo hē* and *lo cantaré* in the 15th century. These results suggest that *cantar lo hē* may not have been the only construction traditionally included under the umbrella of SF that had different functions. I have chosen for this study to maintain the wider definition of the SF construction, and have left the sorting out of potential meaning differences between syntactic options for later analysis. See Section 3.2 for a fuller discussion.
and the PF. This competition is well recognized, as is the fact that the PF appears to be replacing, or, in some American varieties, to have replaced, the SF (Berschin 1986:301; cf. Cartagena 1995-1996): "In colloquial Spanish, …the future tense [SF] is normally used not to express future time…. Future time is signalled by the verb IR (‘to go’) plus the infinitive [PF]" (Palmer 2001:105). The Spanish futures have been studied in various perspectives by a quite few scholars (e.g. Bishop 1973; Claney 1975; Company 1985-1986, 2006; Berschin 1986; Kitova 1986; Sedano 1994; Gutiérrez 1995; Cartagena 1995-1996; Molho 1997; Westmoreland 1997; Company Company & Medina Urrea 1999; Matte Bon 2005, among others). These studies dialogue in interesting and often complementary ways with studies of futures in other Romance languages, most notably French (Wales 1983; Gagnon 1990; Helland 1997; Poplack & Turpin 1999; Larreya 2000; Stage 2002), but also Portuguese (Tlaskal 1978; Poplack & Malvar forthcoming), Italian (Loporcaro 1999; Larreya 2000), Romanian (Ripeanu 1994), and Romance as a whole (Fleishman 1982; Nocentini 2001). These studies are further informed by a few key works on future expression in general (e.g. Ultan 1978; Bybee & Pagliuca 1987; Vet 1994; cf. Pedrero 1993 on periphrastic futures in Indo-European).

1.1.2 Romance futures: two questions

Most of the work on Romance futures focuses on two particular questions (cf. Matte Bon 1995): 1) what differentiates the SF from the PF semantically?; and 2) what meanings does the SF have besides futurity?4 In this section, I will outline the previous research on these two topics.

---

4 Matte Bon links the now widespread practice of identifying semantic nuances of forms to the persistence of the modes of analysis in the tradition of “el estudio de las lenguas clásicas, a su vez marcado por el
1.1.2.1 Semantic differences and speaker choice

When two or more future forms exist at the same time in the same language variety, there are various possibilities for the division of labor of future expression—or for any forms in variation. On one extreme, it could be that the forms always occur in complementary distribution, reflecting inherent semantic differences. On the other (hypothetical) extreme, they could conceivably be in free variation. Most scholars have agreed that the patterns found in speakers' choices regarding PF and SF indicate that Romance futures lie somewhere between these two extreme scenarios, such that speakers may have a choice in some contexts but not in others.

Pinpointing semantic (or other) differences between PF and SF has been no simple task; in fact, it constitutes the single-most recurring question in studies on Romance futures. The linguistic studies that have addressed this question have often focused on the PF, since the SF has traditionally been seen as more "neutral and psychologically detached" than the PF (see Poplack & Turpin 1999:137 for a summary of studies). The perspective that the SF is the Spanish future is perhaps most notable in enfocado didáctico, que llevaba al estudiante a acercarse a estas lenguas a través de la traducción. Esto le empujaba a concebir su descripción intentando detallar los efectos expresivos que podían darse en diferentes contextos” (2005:np).

5 The notion of "free variation," in which two forms are interchangeable and usage is unstructured, is highly problematic, since the basis of variationist studies is the acknowledgement of "structured heterogeneity" (Weinreich, Labov & Herzog 1968:99-100). I include it here only as a conceptual extreme—regardless of its linguistic reality—since some grammarians have, in fact, suggested that the SF and the PF in Spanish are in free variation (Gerboin & Leroy 1991:285; Butt & Benjamin 1994:219).

6 In contrast, Jensen (2002) assigns subjectivity to the SF and objectivity to the PF.
reference grammars, which have often failed to describe the PF as a future form at all (Matte Bon 2005). Cartagena (1995-1996:95) explains this bias in grammars as follows:

"Si la segunda [PF] no ha sido tratada hasta ahora como tiempo verbal, es simplemente porque nuestras gramáticas siguen la tradición gramatical latina, considerando como tiempos verbales sólo a los herederos de los existentes en latín."

The attempt to fill this descriptive gap and to explain this as of yet unacknowledged grammatical phenomenon, despite having had "un desarrollo relativamente lento en el castellano peninsular" (Granda 1997:283), stretches back over five decades (e.g. Kahane & Hutter 1953; Reid 1955; Montes 1962-1963; Imbs 1968; Sáez Godoy 1968; Paufler 1970). Since that time, the PF has been assigned over a dozen meanings, most related to each other. The SF, in turn, has been assigned a similar number of meanings, sometimes defined negatively in opposition to PF, at times with meanings that are also in opposition with each other.  

7 I will not attempt here to provide an exhaustive list of these meanings; I will, however, outline the strongest currents of thought that have dominated hypotheses regarding the choice between PF and SF in Romance future contexts.

---

7 Matte Bon notes, "la mayor parte de los autores intentan clasificar los diferentes usos del futuro [sintético] según los matices que interpretamos en cada contexto específico." He further points out that, "entre los autores que J. Mª Brucart (2000:151-172) clasifica dentro de los los [sic] estudiosos gramaticales de corte tradicional, S. Fernández Ramírez [e.g. 1986] es, sin lugar a dudas, el que presenta la casuística más amplia. ...los demás presentan un número limitado de usos, siguiendo a Bello y a Gili Gaya" (Matte Bon 1995:n4).
The lion's share of hypotheses on semantic differences between PF and SF has centered around three main themes: i) time depth; ii) link to present or moment of speech; and iii) speaker attitude toward (the eventuality of) the event:

i) **Time depth:** In the case of time depth, when cited as the source of difference, the PF is invariably assigned a meaning that places the event closer to the moment of speech. The event may be imminent (e.g. Berschin 1986:303; Gómez Torrego 1988:66; Fernández de Castro 1999:206; Llorente Vigil 1999:25-27; Melis 2006), or just near (e.g. Tlaskal 1978:209; Sedano 1994; Confax 1995); the PF may indicate inchoative aspect\(^8\) (Gili Gaya 1964:107-108; Fente, Fernández & Feijóo 1972:15; Lope Blanch 1983 [1972]; RAE 1973:406; Granda 1997:282) prospective aspect\(^9\) (Cartagena 1995-1996; González Araña and Herrero Aisa 1997:81; Jensen 2002); or it may indicate that the preparatory state necessary for the realization of the event is already achieved (Vet 1993, 1994; Jensen 2002).\(^{10}\) Drawing on its historical source, Bybee and Pagliuca (1987) propose an "agent on a path toward a goal" meaning for PF, which, in fact, encompasses many of the uses mentioned in this paragraph.

---

\(^8\) The term "inchoative aspect" refers to an event that is just beginning to take place. The Real Academia Española aseverates, "**Ir a y echar a forman a menudo expresiones incoativas. 'Ir a + infinitivo' significa acción que comienza a efectuarse, bien en la intención, bien en la realidad objetiva...**," such that the event already has begun to take place (1973:446).

\(^9\) The term "prospective aspect" refers to an event that is about to take place, but has not yet begun. This may or may not imply imminence (Comrie 1976:64).

\(^{10}\) Nevertheless, others claim that the PF is used for general futurity, not just immediate future (e.g. Wales 1983); a few grammarians have suggested that these forms are completely, or nearly completely, interchangeable (Gerboin & Leroy 1991:285; Butt & Benjamin 1994:219).
ii) *Link to present*: The second approach takes into account the subjective interpretation the speaker holds regarding the imminence or nearness of the event. In this perspective, it is not the objective measure of temporal distance that determines variant choice, but rather the speaker's *perception* of that distance. The need for acknowledgement of a speaker's subjective construal of events in the case of futurity is eloquently argued by Gagnon (1990; cf. Gómez Torrego 1988; Matte Bon 2005), though the idea of present relevance, a lack of rupture with the present, or a psychological link with the present are common psychological motives cited for the choice of PF (e.g. Fleischman 1982; Jeanjean 1988; Blanche-Benveniste 1990; Gagnon 1990; Cartagena 1995-1996; Melis 2006; cf. Poplack & Turpin 1999 for a discussion of on psychological motives for variant choice in the literature).

iii) *Speaker attitude*: The third current in semantically based PF-SF distinction draws on the attitude of the speaker alone, with no need for a link to the moment of speech. The psychological meanings cited are nevertheless similar to each other because they are the psychological and/or pragmatic consequences of prospective aspect or completed preparation. Meanings associated with the PF under this approach include certainty (Vet 1994; Confaís 1995; Almeida & Díaz 1998; Jensen 2002), intention (Sedano 1994; Confaís 1995; González Araña and Herrero Aísa 1997:81) or volition and determination (Bishop 1973:89), objectivity (Berschin 1986:303; Jensen 2002), and speaker involvement in the event (Fleischman 1982; Leeman-Bouix 1994), among others. Another meaning that has been associated with the PF is obligation (Llorente Vigil 1999:26).
The SF, on the other hand, has been linked, for instance, with both uncertainty (Imbs 1968; Tlaskal 1978:206-207; Confaís 1995:399) and certainty (Vet 1994); volition (Fernández Ramírez 1986:284-285), determination (Bishop 1973:89), and intention (e.g. Bybee & Pagliuca 1987; Bybee, Pagliuca & Perkins 1991; Villa Crésap 1997:64, 96); nonspecific time (Helland 1997:73); a state in which the preparatory stage is not done (Vet 1994); prospective aspect (Fernández Ramírez 1986:284-285), on the one hand, but rupture with the present (Imbs 1968; Fleischman 1982; Helland 1997; Confaís 1995:398) on the other; absolute future, as a direct expression of speaker attitude (not deictic) (Cartagena 1995-1996); inference and imagination (Reid 1955); politeness and prudence (Imbs 1968); obligation (Bello 1984 [1847]; Fernández Ramírez 1986:284-285; Bybee & Pagliuca 1987); and conjecture (Fernández Ramírez 1986:284-285).

A handful of authors have explored numerous possible semantic differences between Spanish future forms, either through statistical analysis of distribution (e.g. Almeida & Díaz 1998), extensive contemplation and exemplification of the nuances of certain forms in certain contexts (e.g. Vega Llamas 2002; Matte Bon 2005), or both (e.g. Kitova 1986; Sedano 1994; Villa Crésap 1997); in the only experimental study of which I am aware, Berschin (1986) provided stimuli to native speakers in order to discover their preferences for SF or PF in certain contexts. Though most studies have focused on differentiating SF and PF semantically, some have also acknowledged the possibility of neutralization of differences or semantic overlap in some contexts (e.g. Söll 1983:16;

11 Unfortunately, Berschin's (1986) study is relatively short and examines only a handful of contexts. Nevertheless, I call particular attention to his work here because his methodology and results appear to hold much promise.
Wales 1983; Berschin 1986; Vet 1993, 1994; Cartagena 1995-1996; Villa Crésap 1997; Poplack & Turpin 1999; Melis 2006).\textsuperscript{12}

\textbf{1.1.2.2 Epistemic modality and SF semantics}

The second great question in the literature on Spanish and other Romance futures is whether the SF should be considered a tense or mood or both. The difficulty of this question highlights the conceptual (and semantic) complexity involved in attempting to characterize future forms with modal meanings, such as Spanish SF or English \textit{will}, which express both tense (here, referring to events posterior to a moment of reference) and mood (i.e. demonstrating speaker attitude toward the proposition) (Comrie 1985:21, 24).\textsuperscript{13} This is manifested in Spanish SF in particular in epistemic modal uses, as in (1.4), which is not a future, as indicated by the occurrence of \textit{ahora} 'now' and the third-person plural Present of 'be', \textit{están}, in the question. The epistemicity is further emphasized by the preceding \textit{no lo sé si} 'I don't know if'.\textsuperscript{14}

\textsuperscript{12} This approach is not without its nay-sayers; Matte Bon (2005), for instance, explicitly denies the usefulness of the idea of neutralization, perhaps because he (erroneously, in most cases) takes it to mean that researchers believe that variants are interchangeable.

\textsuperscript{13} On epistemic modal uses in Spanish, see: Bello (1984 [1847]:216); Seco (1982 [1954]:76); Montes (1962-1963); Sáez Godoy (1968); Avila (1968); Söll (1968); Valdez (1969); Iuliano (1976); Moreno de Alba (1977); Iuliano and Stefano (1979); Gómez Manzano (1988); Bauhr (1989); Silva-Corvalán and Terrell (1989); Sánchez and Ferrer (1990); Alarcos Llorach (1994:155); Matte Bon (2005); cf. Azevedo (1992:116); Sedano (1994:226); Villa-Crésap (1997:58); Jensen (2002).

\textsuperscript{14} Scholars in Romance linguistics have varied in the way they have referred to this use. For example, Tomaszkiewicz (1988) and Sedano (1994) call it simply the "modal" use, while Cartagena (1995-1996) calls it "probabilistic," Pedretti (1999) "hypothetical," and Butt and Benjamin (1994) "suppositional." For
(1.4) 
- ¿Y Paquito y María dónde están ahora?
- No lo sé si estarán con sus abuelos o estarán en Valencia dónde andarán.
(COREC, CCCON019A, 20s)

'- And Paquito and María, where are they now?
- I don't know if they might be (SF) with their grandparents or they might be (SF) in Valencia where they might be (SF).'

The question of the most accurate way to describe forms with both future and non-future modal meanings, like the SF, as a mood with some temporal uses (e.g. Sweetser 1982; Traugott 1989) or both tense and mood (e.g. Coates 1983; Palmer 2001), is still unresolved (Heine 1995:18 and throughout). Nevertheless, it is often desirable to characterize polysemous forms, if polysemy can be identified, in such a way that occurrences of different meanings may be isolated for study. This can be especially vital in studies of variation (Wolfram 1997:195). It is this desire that has motivated many of the studies of the semantics of the Spanish SF and its counterparts in other languages.

In addressing whether the use of Future forms in both future temporal reference and in non-future epistemic contexts should be considered to represent the use of one gram or two grams in semantic analyses, Comrie (1985:47) states, "we will suggest that there is indeed need to refer independently to future time reference, but it is of course possible that subsequent work may lead to revision of this claim, by showing that these instances of apparent relevance of future time reference can in fact be subsumed under some more general non-tense category." In Chapters 3 and 6 I will present quantitative evidence that Comrie's (1985:47) suggestion to separate future temporal reference and non-future temporal reference occurrences of future markers, like the SF, during the clarity here, to distinguish this use from agent-oriented modal uses of the SF, I will refer to this as "epistemic" SF throughout.
process of description and analysis is felicitous in the case of Peninsular Spanish SF, since the distributional patterns of temporal and epistemic SF demonstrate statistically significant differences (see Sections 3.5 and 6.3).

From a synchronic perspective, some scholars have proposed the need for a continuum-like vision of temporal and non-temporal epistemic meanings in futures (but see Pedretti 1999). A particularly insightful work on the problematic nature of the idea of categorization of epistemic and temporal occurrences of constructions that can occur in contexts with future temporal reference as two distinct grams is the study of the Polish Future by Vettet and Skibinska (1998), which includes a well-articulated defense of the notion of gradient categories, integrating qualitative diachronic and cross-linguistic (Polish and French) evidence into an explanation and critique of earlier accounts of the epistemic/temporal distinction in Future forms. Similarly, in a study of examples drawn from Latin texts, Mellet (1989:277) notes the pragmatic role of grammatical person, along with temporal reference, in determining the relative level of epistemicity in any given occurrence. Though Mellet (1989:276) presents a two-part model for determining how epistemic or temporal an occurrence is, this model is not meant to definitively delimit two separate categories or meanings:

Le futur prouve à l'évidence qu'il faut établir un continuum entre ces deux catégories grammaticales: on ne peut pas se contenter d'opposer deux à deux les différentes formes d'un système en les enfermant dans les cases rigoureusement étanches d'un tableau d'oppositions binaires, comme le font encore trop souvent des analyses guidées par un structuralisme étroit et soumis au modèle de la phonologie [The future proves that one must establish a continuum between these two grammatical categories: one cannot be content to juxtapose the different forms in a system two-by-two by enclosing them in rigourously airtight cells of a table of binary oppositions, as analyses guided by a strict structuralism under the phonological model too often still do] (Mellet 1989:277).
The insistence on a continuum model for temporal-epistemic readings of the SF is likely a reflection of the historical fact that the epistemic use of the SF is a more grammaticized (i.e. further developed and semantically generalized) use that has its origins in the obligation source of the temporal use (Bybee, Pagliuca & Perkins 1991), which in fact is a manifestation of a general cross-linguistic tendency of future forms to take on modal functions (Heine 1995:17-18). While the diachronic relationship between the temporal and epistemic uses of the SF is an excellent explanation for the difficulties scholars have encountered, this knowledge only highlights further the inadequacy of traditional notion of (discrete) categories (see Taylor 1989 for a discussion of the notion of category; see also Thompson & Hopper 2001; and Thompson 2002 on complement clauses). A handful of studies have presented quantitative information regarding epistemic uses of Romance futures: Durán Urrea and Gradoville (2006) report a rate of 79% (N=60/76) of epistemic meaning in SF in spoken New Mexican Spanish, and Villa-Crésap (1997:58) found a rate of 70% from New Mexican spoken Spanish; similarly, Sedano (1994:231) found that epistemic uses of SF made up 58% (N=148/249) of all SF uses in her spoken data from Venezuela; in contrast, Cartagena (1995-1996) found that they made up an average of 11% (N=120/1133) in cross-dialectal study of literary texts written between 1964 and 1982 (see also Bybee & Fleischman 1995 and works therein on modality).¹⁵

Perhaps the general lack of quantitative studies of this use is due to the debate about how to distinguish such uses from "future" uses, which has yet to be satisfactorily

---

¹⁵ Cartagena (1995-1996:83-84) notes, however, that this rate varies greatly by author, from 0% to 56%.
resolved (see Heine 1995; and Narrog 2005 for discussions of this debate). In the present study, like Sedano (1994), I have attempted to distinguish these uses as required for statistical analyses (following Poplack & Turpin 1999; Poplack & Malvar forthcoming), based on the relatively objective criterion of the possibility of future temporal reference, thus defining "epistemic" uses negatively, as those that are not future. This makeshift solution, while revealing distributional tendencies that lend support to its validity (see Chapters 3 and 6), and thereby lending support to Comrie's (1985:47) distinction, does not definitively resolve the ongoing debate about the relationship between futurity and epistemicity or how to classify specific occurrences of these forms.

The present study is a quantitative approach to both the question of variation in form (choice of PF or SF) and the question of variation in function (polysemy within PF and especially SF) detailed above. I have drawn on findings in previous studies, with particular attention to the diachronic and quantitative approaches (e.g. Company 1985-1986; Kitova 1986; Bybee & Pagliuca 1987; Westmoreland 1997; Poplack & Turpin 1999; Poplack & Malvar forthcoming). In this case, such studies are particularly valuable, for they offer suggestions for empirical observation of Spanish futures within a usage-based framework, as outlined in Section 1.2.

1.2 A usage-based approach

Since the 1970s, the field of linguistics has benefited from a new approach to grammar in which the structure of language is seen as inseparable from its function as a communicative social tool (e.g. Bolinger 1976; Givón 1979; Hopper 1987, 1991; 16 Both Poplack and Tagliamonte (2001) and Poplack and Malvar (forthcoming) excluded non-future epistemic uses from their analyses; neither study provides the token count of epistemic uses excluded.
This approach, called the "functional" approach, asks fundamentally different questions about language than the "generative" approach that gained momentum in the 1960s through the work of scholars such as Noam Chomsky (e.g. 1957, 1965); while the generativists were concerned about language "as an innate capacity of the individual," functionalists "[privilege] language as a device for communication between speakers and addressees" (Brinton & Traugott 2005:3). In a generative approach to language, what is interesting is the nature and origins of the system of knowledge. Brinton and Traugott (2005:3) describe the generative approach as follows:

---

17 Of course, this idea was not completely new. Historical linguists had long acknowledged, at times only implicitly, the importance of use in language change (cf. especially Meillet 1912/1965). Company (2003:5) notes, "la gramática histórica siempre operó con un concepto muy amplio de significado, que incluía tanto semántica léxica, cuanto la conceptualización subjetiva del hablante, así como el contexto extralingüístico, la capacidad de los signos de ser empleados icónicamente, además de cuestiones pragmáticas y estilísticas, para poder dar cuenta cabalmente de un aspecto esencial de los signos, a saber, el dinamismo de la relación entre significado y significante y del signo con lo comunicado, así como del hecho esencial de que los significados no son estables ni indisolubles de las formas y que cualquier proceso de cambio supone una dinámica de pérdida y ganancia semántica, en donde las formas entran a nuevos contextos recargándose de nuevos significados y haciéndose cada vez más polisémicas, a la vez que significados viejos pueden debilitarse e incluso llegar a desaparecer." See also William Labov's (e.g. 1966/1982, 1969, 1971, 1972, 1976, etc.) early work on contextual social and linguistic factors that conditioned variation and change in the speech communities he studied, leading to the blossoming of the empirically-based field of sociolinguistics.

The assumption is that the language capacity is computational and syntactic, and by hypothesis optimally structured and ultimately binary in nature. It is a self-contained modular mechanism that does not reflect external factors such as cultural or social systems. Nor does it reflect experiential structures such as vision or production factors such as frequency (this is known as the hypothesis of 'autonomous syntax'). The universals of language that are posited are absolute in the sense that one counterexample disproves them (see, e.g., Newmeyer 1998:263).

Here, then, "the role of experience in learning a language involves little more than the setting of parameters" (Langacker 2000:2).

About a decade after the advent of generative approach, some linguists began to ask "how speakers can use the 'bricolage' or 'jerry-built structure[s]' (Bolinger 1976:1) of language to impart information, and to get things done" (Brinton & Traugott 2005:3; cf. Hopper 1987). In the same spirit, in 1987, cognitive linguist Ronald Langacker proposed a model of language that aimed for psychological reality, a model that he termed "usage-based" (Langacker 1987a:492), on the grounds that

it cannot be maintained on purely methodological grounds that the most parsimonious grammar is the best one. Should it prove that the cognitive representation of language is in fact massive and highly redundant, the most accurate description of it (as a psychological entity) will reflect that size and redundancy (Langacker 2000:2).

This "non-reductive," "maximalist," "bottom-up" (Langacker 1988) model is one in which "all facets of linguistic structure can be reasonably described" (Langacker 2000).

In this perspective, the object of interest is language use, both in language-specific grammar and cross-linguistic comparison (Brinton & Traugott 2005:3-4).

One far-reaching implication of a usage-based perspective, of keen interest for the present study, is that our concept of grammatical category must reflect actual use. Thus, in a functionalist perspective, linguistic categories are not discrete or homogenous, nor are they defined by necessary and sufficient conditions; instead, categories are "espacios
no discretos, sino inestables, flexibles, redefinibles y manipulables de manera creativa por los hablantes" (Company 2003:5; see also Hopper & Thompson 1984; Thompson & Hopper 2001). Both within the category and between categories, then, there is a continuum of linguistic forms, with various focal points, "donde se sitúan las entradas léxicas que son mejores representantes de la categoría, el prototipo, y límites categoriales no nítidos ni bien establecidos, e incluso algunas entradas léxicas pueden estar situadas en zonas fronterizas y exhibir las propiedades de dos o más categorías" (Company 2003:6). For example, count nouns like chair or apple are better examples of the category "noun" than are, say, abstract nouns such as love or responsibility, since the prototypical noun has the characteristics of being concrete, visible, tangible, delimitable and divisible into parts (Langacker 1987b:58-60; Company 2001, 2003). In other words, it is no longer simply a question of whether a lexical item is an example of A or B, but rather of how prototypical a representation is it of A or B; it may even be a peripheral member of both A and B (Givón 1986; Lakoff 1987:esp. Ch. 1; Seiler 1989; Taylor 1989; Company 1997, 2003; Thompson & Hopper 2001; Thompson 2002).

In taking the basic perspective that language use and language structure are inseparable, we also commit ourselves to a certain kind of data as the basis of study: actual use. Thus, like most other studies done within a usage-based framework, the present study is text-based, i.e., it takes as the basis of analysis the actual occurrences of forms in texts and speech.19 Along with the acceptance of language use as giving rise to

---

19 One exception to this reliance on solely naturally occurring data is the case of wide-scale cross-linguistic study, such as Bybee, Perkins & Pagliuca (1994), which examines patterns in 76 world languages, many of which have no natural data recorded or available for analysis; instead, for this and similar projects, scholars have based their analyses on information gleaned from reference grammars, which aim to describe
language structure comes the simultaneous recognition that linguistic structure is never fixed, and that synchrony is simply a snapshot of diachrony (Weinreich, Labov & Herzog 1968; Company 2003:25). One of the most intriguing aspects of this approach is the perspective it offers for studying (and postulating) language change. My interest lies specifically with how speakers create and transform grammar when negotiating future expression.

1.2.1 Grammaticization

Grammaticization theory, a model of the creation of grammar (cf. Hopper 1987) within a usage-based framework (e.g. Lehmann 1982; Hopper & Traugott 1993; Bybee, Perkins & Pagliuca 1994), is a theory about how grammar is made and why it changes in the ways that it does. The development of this framework has been little less than a revolution in the way linguists think about and describe grammar. Such a preoccupation is not new; in fact, the idea of what we now call grammaticization has been described and discussed since the 18th century (cf. Lehmann 1982:Ch. 1; Heine, Claudi & Hünnemeyer 1991:Ch. 1; Hopper & Traugott 1993:Ch. 2; Heine 2003; Thompson 2003). The term "grammaticalisation" (used interchangeably with "grammaticization") was not coined until the early 20th century, by Meillet (1912/1965:131), who defined it as "...le passage d'un mot autonome au rôle d'élément grammatical...l'attribution du caractère grammatical as it is used. As Bybee and Pagliuca (1987:110) explain, "The languages of Perkins' sample were randomly chosen, controlling for genetic and areal bias. If one wants to arrive at valid statements concerning the relative frequency of a given linguistic phenomenon in the languages of the world, then only samples so constructed may be used; convenience samples, because they lack the controls, are inappropriate."
à un mot jadis autonome [...the movement of an autonomous word into the role of grammatical element...the attribution of grammatical character to an erstwhile autonomous word]." While Meillet's characterization of grammaticization appears to be concerned only with lexical items becoming grammatical items, he also noted that word order changes, such as a more restricted word order in Romance languages than in Latin, may also be considered cases of grammaticization (Traugott 2003:624-625).

The modern definition of grammaticization has not changed in this basic notion, though recent literature (e.g. Traugott 2003) has emphasized the important role of constructions in the process (see note 21). Company (2003:9) offers the following basic definition: "se trata de un proceso mediante el cual una forma o construcción asume una función gramatical, o bien una entidad o construcción ya gramatical adquiere una función aún más gramatical.' This process is irreversible, unidirectional and gradual (Heine & Reh 1984:74-76; Givón 1979; Lehmann 1982; Bybee, Perkins & Pagliuca 1994:12-14), and has been characterized as a movement towards abstraction (McMahon 1994, Ch. 2).

In this study, I view language change as a gradual process that takes place through what Langacker (2000:3) refers to as "entrenchment," which has also been termed "habit formation," "automatization," and "routinization" (Haiman 1994). The idea is that:

The occurrence of psychological events leaves some kind of trace that facilitates their re-occurrence. Through repetition, even a highly complex event can coalesce into a well-rehearsed routine that is easily elicited and reliably executed. When a complex structure comes to be manipulable as a 'pre-packaged' assembly, no longer requiring conscious attention to its

---

20 One of the major concerns of grammaticization theory is mechanisms of change. While there has been much debate surrounding this question, I do not intend to enter into it here (see, e.g., Sweetser 1988 on metaphor as the basic mechanism for change; and Haspelmath 1998 for a critique of the notion of reanalysis in grammaticization).
parts or their arrangement, I say that it has the status of a unit (Langacker 2000:3-4, emphasis in original; cf. Hopper 1987, 1998; Company 2002a).

Given such a process, how do speakers choose which events to routinize, and what does routinization look like? One theoretical breakthrough has been the discovery—based on empirical cross-linguistic evidence—that there are "major paths of diachronic development" (Bybee, Perkins & Pagliuca 1994:3; cf. Bybee & Dahl 1989), and that these are recurrent because they rely on the conventionalization of commonly occurring pragmatic-discursive strategies (Traugott 1982, 1989; Givon 1984; Girón Alconchel 2002:104-105) as a mechanism for change. The sources for grammar are "basic to human experience and largely culturally independent" (Bybee, Perkins & Pagliuca 1994:10) and thus "tend to be conceived of in a similar way across linguistic and ethnic boundaries" (Heine, Claudi & Hün nemeyer 1991:33). For example, one common path is for verbs of possession, in certain constructions, to develop, or "grammaticize," into verbs of obligation, such as English have to (Bybee, Perkins & Pagliuca 1994:7-8) or Spanish tener que 'have to'.\footnote{As Lehmann (1992:406) notes, "grammaticalization does not merely seize a word or morpheme...but the whole construction formed by the syntagmatic relations of the element in question." See Traugott (2003) for a thorough discussion of the role of constructions in grammatization (cf. Bybee 2003:146). See Bybee, Perkins and Pagliuca (1994:168) on the importance of context and construction for English be going to.} Company (2003:21) offers the following model for the underlying tendencies of language change that compete in the realization of grammaticization.
Figure 1.1. *Force dynamic in the generation of a grammaticization* (from *Company 2003:12*)

<table>
<thead>
<tr>
<th>Conservative tendency</th>
<th>Innovative tendency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbolic, routinized, automatized signs</td>
<td>Highly iconic or transparent signs</td>
</tr>
<tr>
<td>Articulatory separation, perceptive ease</td>
<td>Less effort, speed, erosion</td>
</tr>
<tr>
<td>Isomorphism (1 form = 1 meaning)</td>
<td>Economy</td>
</tr>
<tr>
<td>Pressure to use standard morpho-syntax</td>
<td>Discursive-pragmatic pressure</td>
</tr>
<tr>
<td>Referential descriptive pressure</td>
<td>Evaluative, subjective pressure</td>
</tr>
<tr>
<td>Analogic pressure</td>
<td>Pressure towards irregularity</td>
</tr>
<tr>
<td>Pressure to identify with a group</td>
<td>Pressure to distinguish oneself from the group</td>
</tr>
</tbody>
</table>

While grammaticization is a common universal phenomenon (Bybee, Perkins & Pagliuca 1994), not just any kind of change is subsumed under grammaticization. "Grammaticization," as Traugott (2003:645) notes, "is not coterminous with morphosyntactic change," but rather "the process whereby lexical material in highly constrained pragmatic and morphosyntactic contexts is assigned grammatical function, and once grammatical, is assigned increasingly grammatical, operator-like function." In this vision of the creation of grammar, change is gradual, gradient, continuous, context-based and usage-dependent.

A key advantage of grammaticization theory is that it allows us to make predictions about the diachronic trajectory of forms undergoing change. If we know the lexical source of a particular morpheme, we can make predictions about its semantic development over time. In turn, if we are interested in a particular semantic space, such as future, the paths proposed, for example, by Bybee and Dahl (1989), would guide us in a search for the (lexical) origins of forms that may serve the function of interest. In the case of corpus studies of diachronic data, the proposed paths of grammaticization may help to make sense of the gradual shifts in a form's semantics and distributional patterns.

1.2.2 Linguistic correlates of grammaticization

The process of grammaticization involves, of course, semantic change, but there are also other far-reaching co-occurring phenomena, which are highly constrained and interrelated. Some of the most important transformations seen in a construction undergoing grammaticization include:

i) In the beginning, a rise in frequency (cf. Bybee, Perkins & Pagliuca 1994:8);

ii) Change in category membership (Company 2003);

iii) Loss of syntactic freedom (Hopper and Traugott 1993:122-123); and

iv) Loss of components of meaning, leading to semantic generalization, through a process of "semantic bleaching" (e.g. Heine & Reh 1984:15);

v) Functional generalization, such that the construction becomes ever more useful (extending to new functions, Bybee, Perkins & Pagliuca 1994:6).

A model of gradual syntactic change that relies upon usage also has profound consequences for our vision of the "structure" of language. If forms are slowly moving in and out of use, often in the same semantic spaces, there will be times when forms compete for this space. This situation has been key in the work of sociolinguists for decades, and refers to the inherent structured variability in language (cf. Weinreich, Labov & Herzog 1968:99-100). Another possibility in grammaticization is that, as a gram
changes semantically, older uses may not disappear, but rather co-exist, perhaps indefinitely, a phenomenon that Bybee and Pagliuca (1987) called "retention" and Hopper (1991) called "persistence." This situation may lead to divergence of forms (Hopper 1991). It also may show itself in the distributional patterns of grammaticizing forms, as in Torres Cacoullos (1999:29-34), where a detailed quantitative case is made for retention reflected in distribution constraints.

1.2.3 Methods in a usage-based framework

If linguistic variation is accepted as a natural and inherent part of language, this must necessarily inform the methods used to analyze and interpret linguistic phenomena. One implication is that the positing of "acceptable" and "unacceptable" contexts for constructions (illustrated with constructed examples) is no longer interesting; instead, it

---

22 The term "gram" refers to grammatical morphemes, as opposed to lexical morphemes; the former commonly develop from the latter (see Bybee & Dahl 1989).

23 Unfortunately, the findings in studies that rely on constructed examples cannot be verified without empirical study. The use of constructed examples in most cases is necessary to the line of argument, which posits contexts in which the forms in question are or are not acceptable, demonstrated by examples of "unacceptable" or unlikely phrases, which, of course, if properly identified, should not appear in natural data. Such judgments are based on the intuitions of the researchers, which are not necessarily accurate. In the case of Canadian French, for example, Poplack and Turpin (1999:140-142) found that postulated non-existent utterances occurred frequently in their corpus. For example, in their examination of the use of the verbs se marier 'get married' in Ottawa-Hull spoken French, they found that 55% of the uses of se marier in contexts in which the speaker was uncertain of the realization of the event occurred with PF; other scholars (e.g. Imbs 1968; Fleischman 1982; Confais 1995) had deemed this context unsuitable for PF use, and used this as evidence for the "undeniable" difference in meaning between Paul va se marier 'Paul is going to get
is often not a question of possible or impossible, but of likely or unlikely. As D. Sankoff
notes, "When scientifically accounting for an entire speech sample or corpus, striking and
widespread regularities may emerge which pertain solely to the relative frequency of
occurrence of various structures, rather than to their existence or grammaticality"
(1988a:141). Given this perspective, scholars acknowledging variation want to know how
likely or unlikely it is, for example, that a certain form or construction will be chosen in a
certain linguistic context or by certain speakers in certain social circumstances. Such
issues, which involve relative tendencies and not absolute rules, benefit from the use of
quantitative analysis (see García 1985b for a discussion of the theoretical implications of
the relation of quantitative data to grammar).

One approach that has shown itself to be particularly successful in using
quantitative analyses in the discovery of grammaticization in progress in natural language
in use, is variation theory (e.g. D. Sankoff 1988a; G. Sankoff 1974, 1990; Poplack &
Turpin 1999; Torres Cacoullos 2000, 2001; Poplack & Tagliamonte 2001; Schwenter &
Torres Cacoullos 2005; cf. Bayley 2002). Variation theory took its inspiration from the
quantitative tradition of sociolinguistics and the foundational work of William Labov
married (PF) and Paul se mariera 'Paul will marry (SF)' (Poplack & Turpin 1999:140). As Poplack and
Malvar (forthcoming) incisively note,

...very few of the motivations ascribed to variant choice in the (prescriptive or
descriptive) literature are now, or ever have been, relevant to actual usage. Grammarians
have been silent on the role of the operative contextual factors, focusing instead on
semantic, psychological and other motivations which have no basis in empirical fact.

It is encouraging to note that an operationalization of the semantic and psychological motivations
highlighted in works that rely on constructed examples—where methodologically possible—does lead to
the discovery of significant distributional differences among the Spanish PF and SF (see Chapter 4), though
not always in the directions predicted (cf. Poplack & Malvar forthcoming).
(e.g. 1966/1982, 1969, 1971, 1982, 1984; etc.; cf. G. Sankoff 1974; D. Sankoff 1988a; Chambers 1995; Poplack and Tagliamonte 2001:7). A key component of variationist work is that it is quantitative, focusing not on the grammaticality or the existence of structures, but rather on relative frequencies of structures (D. Sankoff 1988a:141).

Grammaticization theory clearly predicts variation (and even earlier scholars recognized that "all change involves variability and heterogeneity" (Weinreich, Labov & Herzog 1968:188)) and offers a framework for explaining in diachronic terms any variation that may be found within one construction (i.e. the development of a construction). As Bybee, Perkins & Pagliuca (1994:363) assert, the analysis of "the mechanisms of change that create grammatical meaning must proceed from analysis of the use of grams as these changes are taking place." Without this crucial step, we have a prediction and an explanation, but no local process, leaving us with no grounds to assert the accuracy of general predictions in the complex manifestation of local phenomena. One way to discover changes as they are taking place is through quantitative study of variation between constructions for without it, "the most that could be said is that the language admitted two (or more) options" (Chambers 1995:205).

This study focuses on variation and change in the two Spanish futures. The objectives are to determine, i) how speakers' choices regarding the future forms SF and PF changed during the process of grammaticization and ii) the effect, if any, of

---

24 For use of quantitative methods in non-variationist studies of variation within a grammaticization framework, see, e.g., Thompson & Mulac 1991a, 1991b; Coates 1995; Heine 1995.

25 This approach is in line with Weinreich, Labov and Herzog's suggestion: "We feel it important to dwell explicitly on empirical foundations, in view of the conscious or unconscious disregard of empirical principles which pervades some of the most influential work in linguistics today" (1968:103).
(changing) token frequency on these patterns. The most appropriate methodology for going about answering these questions is the quantitative variationist methodology utilized within the framework of grammaticization theory. An amalgamation of variationist methodology and grammaticization theory offers the ideal combination of an empirically sound and replicable scientific methodology and a theory about well-attested cross-linguistic tendencies within a model of language based on natural language use. This work will add to the growing body of literature in usage-based variationist studies within a grammaticization framework (see, e.g., Poplack & Turpin 1999; Torres Cacoullos 2000, 2001; Poplack & Tagliamonte 2001; Schwenter & Torres Cacoullos 2005; Walker & Torres Cacoullos 2005; Poplack & Malvar forthcoming).

While grammaticization theory and variation theory stem from different traditions, some pioneering scholars were at once aware of the tremendous advantages offered by the application of variationist methodology to the theoretical issues raised in the study of grammaticization (e.g., G. Sankoff 1974; Ash & Myhill 1986; Myhill 1992). The insight provided by grammaticization theory in the explanation of variation is twofold: i) forms tend to follow universal paths of grammaticization, such that certain forms with certain meanings are the most likely candidates to take certain new meanings through grammaticization (e.g. resultative > anterior > perfective or past (Bybee, Perkins & Pagliuca 1994:68, 105)); and ii) retention of older meanings may constrain the new meanings a form may take. This study, then, will aim both at accurate description of variation and cross-linguistically informed explanation of the local manifestations of apparently universal phenomena.
1.3 Variation theory

1.3.1 Structured variability and competition

Variation theory, and the quantitative paradigm on which it is founded (Bayley 2002), offers one solution to the problem of discovering the systematic patterns in dynamic processes of change. Variationists acknowledge inherent variability in language. Furthermore, variationists also recognize that variation patterns are not random, but rather are characterized by "structured heterogeneity" (Weinreich, Labov & Herzog 1968:99-100; cf. Bayley 2002:117, 122). These assumptions are in sharp contrast to those held by linguists in the generative and other traditions, who have often relied on constructed or "cleaned-up" data, and in doing so have been able to ignore variation in formulating theories (see Milroy & Milroy 1997:48 for a discussion of these methodological problems).

The basic unit of study in variationist studies is the linguistic variable. A variable in variation studies is an abstraction "made up of a class of variants—varying items that exist in a structurally-defined set of some type" (Wolfram 1997:195, emphasis in original). In the context of the discussion above, then, variation studies are preoccupied primarily with the kind of synchronic variation predicted in grammaticization theory that I have been referring to as variation in form, or "layering" (Hopper 2001). In the variationist framework, the variation among forms or constructions is understood to be conditioned by various probabilistic constraints. These constraints, then, are not rules.

26 In fact, Weinreich, Labov and Herzog assert that, "in a language serving a complex (i.e., real) community, it is absence of structured heterogeneity that would be dysfunctional" (1968:101, emphasis in original).
but rather constraints on the optionality of forms (see Langacker 2000:3 on the notion of rules in generativist and functionalist models of language), with some constraints being linguistic in nature (e.g. Poplack & Tagliamonte 2001), some sociological (e.g. Wolfram 1997:197), and others stylistic (e.g. Silva-Corvalán 1997:115). For example, the use of the PF in Spanish has been hypothesized to be constrained by the linguistic factor of co-occurrence with a temporal adverbial (e.g. Cartagena 1978:384; Fleischman 1982:87, 181; Helland 1997:73-75; Melis 2006), by the social factor of the socio-economic class of the speaker (Gili Gaya 1964:45; Cartagena 1995-1996:80; Aaron 2003; but see Poplack & Turpin 1999:161), and by the stylistic factor of the level of formality in the interactional situation (Alvarez & Barros 1981:8; Hernández Alonso 1968:32; Silva-Corvalán & Terrell 1989:206; Sedano 1994:226; cf. Vet 1994:53 on hypotheses about PF-SF variation that posit stylistic differences). The goal of variation studies, then, is to discover these constraints.

A crucial component of variationist studies is that they are guided by Labov's (1972) principle of accountability: the discovery of constraints is only possible if all tokens of a particular structure or conceptual domain are analyzed "by tabulating the number of actual occurrences of a particular structure in terms of all those cases where a form might have occurred, or potential cases" (Wolfram 1997:206, emphasis in original; cf. Bayley 2002:123). It is only through the examination of all potential occurrences, then, that we may reliably account for the structured variation found in natural language. Ideally, then, in the case of future expression, for example, this would involve the examination of all future contexts, regardless of the form that is used. This methodology is essential, since only here will we discover not only when a form can occur, but also
when it *cannot*, and, of more interest, when it *tends* to occur and when it *tends not* to occur, that is, its likelihood of co-occurrence with the contextual features examined (Bayley 2002:118). The present study is limited to only two constructions, the PF and the SF.

### 1.3.2 Frequency and language change

One of the most important components of changes in progress is increased frequency (e.g. Bybee & Hopper 2001), either as a possible indicator of change (e.g. Labov 1994:43-63) or as a contributing factor in speeding up the processes of grammaticization, including automatization of production and "chunking," semantic generalization, and phonological reduction (Bybee & Hopper 2001; cf. Langacker 2000 on "entrenchment").

An important question that is just beginning to be addressed is the role of frequency in conditioning on variant choice when two or more forms compete within the same semantic domain (Poplack & Malvar forthcoming; on competition in grammaticization, see Bybee, Perkins & Pagliuca 1994:232-235).

Poplack (2001), in examining possible frequency effects on variation patterns in three areas of the "irrealis" domain in French, found conflicting results, indicating the need for more studies in order to define the scope of frequency effects. For example, how does the increased frequency of a new form affect the semantic domains of existing forms with overlapping uses? As forms develop side-by-side, how do the linguistic constraints on their use change as their relative rates increase or decrease in their overlapping space?

---

27 The body of literature on frequency is rich; a review of this literature, however, is beyond the scope of this study. Readers interested in this topic are encouraged to examine especially Bybee and Hopper (2001) and, for phonology, Bybee (2001b).
One answer is Kroch's (1989) Constant Rate Hypothesis, which predicts that rates of change and linguistic constraints are independent of one another, such that, as one variant replaces another, it does so at the same rate across all contexts. Poplack and Malvar suggest, however, that "the applicability of the Constant Rate Effect must be restricted to periods in which the repertoire of variants remains stable, i.e. periods which feature neither incipient nor moribund forms" (forthcoming), a suggestion which follows from their finding that constraints are transferred from one variant to another as variants enter and exit the variable context. Can the development of an incipient form "push" older forms into semantic or syntactic spaces in which competition is less intense or non-existent (e.g. old future forms possibly becoming subjunctives, suggested in Bybee, Perkins & Pagliuca 1994:235)?

Stein (1995), in his study on the subjectification of inverted word order in English, notes that a construction may be able to conventionalize certain pragmatic/subjective meanings only when it is in variation with a less-marked alternative. Company Company and Medina Urrea (1999) provide support for this hypothesis in their analysis of the rise of the synthetic form of SF (e.g. lo cantaré) at the expense of its analytic counterpart (e.g. cantar lo hé) in Old Spanish.

In a variationist study of the diachronic changes in constraints on four variants of Brazilian Portuguese, Poplack and Malvar (forthcoming) provide evidence that receding constructions are particularly vulnerable to functional changes: "As the older variants recede, they lose or transfer their associated constraints to incoming variants, with the striking result that the structure of the future domain remains the same, though its functions are redistributed. Thus the important distinctions (functional, semantic)
continue to be expressed, albeit by different exponents" (cf. Bybee & Thompson 2000 and Bybee & Hopper 2001 on the role of frequency in syntax and linguistic structure, respectively). This is evidence that the relative frequency of variants can affect the patterning of constraints on variation. I aim to contribute to this dialog on frequency and variation in Chapter 6, where I present evidence that the generalization of the PF into futurity may have triggered the move of the SF into epistemic modal meaning, which, in turn, affected 20th-century PF-SF variation patterns.

1.4 The present study

The Spanish and other Romance future tenses have been the object of many years of linguistic study. Despite this attention, only a handful of studies have examined the influence of multiple factors that may condition variation between competing Spanish future forms in natural data (e.g. Sedano 1994; Cartagena 1995-1996; Villa Crésap 1997; Almeida & Díaz 1998; also Company and Medina Urrea 1999 on analytic vs. synthetic SF). Except in the case of Villa Crésap (1997) and Company and Medina Urrea (1999), however, these studies are small in scope and do not include diachronic data. One of the reasons that so few studies have been able to explain how speakers choose one form over the other is that the question of how future forms are actually used in texts or discourse is not examined at all (e.g. Helland 1997; cf. Poplack & Malvar forthcoming). Niño-Murcia (1992:708), for example, decries the fact that Keniston's (1937) and Kany's (1969 [1963]) early work on Spanish future expression do not take into account semantic complexities, an omission which she attributes to "la tendencia de estos autores a citar los futuros totalmente fuera de contexto" (see Vega Llamas 2002:6-7, for a similar complaint).
Other scholars, while using natural, contextualized data, may list meanings without providing an overall idea of frequency (e.g. Fernández Ramírez 1986), or they may present data for only one or two factors hypothesized to condition variation, such as imminence (e.g. Melis 2006). Matte Bon (2005, emphasis added), in reference to Fernández Ramírez' (1986) treatment of the SF, explains what is lacking in such studies:

Más allá de la admiración que causan en nosotros el honesto y encomiable esfuerzo por clasificar de manera rigurosa una gama muy variada de efectos expresivos sin perder de vista ningún matiz, y la ingente labor de observación sistemática de la realidad de la lengua en contextos naturales emprendida por este autor [Fernández Ramírez], el enfoque adoptado suscita cierta perplejidad porque, por más que nos identifiquemos con sus observaciones en buena parte de los usos registrados, se tiene la sensación de un trabajo demasiado anclado en una tradición de estudios gramaticales que enumeraba usos, y aun cuando busca el elemento en común que los une a todos, no consigue dar cuenta del sistema esencial que es la lengua debido al peso excesivo que atribuye a cada ejemplo y cada interpretación.

This study aims to address these missing links through the combination of a usage-based approach to grammaticization (e.g. Bybee, Perkins & Pagliuca 1994) and variationist methodology (e.g. Poplack and Tagliamonte 2001).

This study consists of an explanation of the data and methods (Chapter 2), a (form-based) analysis of the evolution of the SF (Chapter 3) and of the PF (Chapter 4), and a diachronic quantitative analysis of the functional space in which these two forms have overlapped (Chapter 5). Chapter 6 exploits the benefits of combining form- and function-based analyses in order to more fully comprehend the nature of language change when forms compete while both are grammaticizing. A summary of my findings, limitations of this study, and suggestions for further research are offered in Chapter 7.
2 Methodology

2.1 Corpus

The corpus used in this study was comprised of 17 documents representing the mid-13th through the early 21st centuries, and included 16 texts from the written genre and one transcription of 20th-century spoken data. The documents chosen for this study, all produced in Spain, were selected in diachronic increments of approximately 150-200 years. This time period was considered sufficient to observe change in progress in written language; a lapse of only 50 years would most likely be too short to note any significant systematic shift in patterns, while a lapse of 300 years would most likely leave intermediate steps in change out of the picture. In the majority of cases, these particular texts were also chosen because they are conservative critical editions and thus most faithful to the original texts. The spoken corpus used is the conversational section from a much longer transcription of 20th-century spoken Peninsular Spanish, Corpus de Referencia de la Lengua Española Contemporánea: Corpus Oral Peninsular (directed by Francisco Marcos Marín; see http://www.illf.uam.es/~fmarcos/informes/corpus/corpulee.html for details), which is available electronically. The total word count for the corpus is approximately 935,500 words, with the following breakdown: Old Spanish (13th-15th centuries), 311,000; 17th century, 232,000; 19th century, 71,500; 20th-century written, 79,000; 20th-century spoken, 242,000. See the list of texts, divided by century/data set, in the Appendix.28

28 For some analyses, data from all Old Spanish (13th-15th centuries) texts are combined due to low token frequencies of the PF and epistemic SF in this time period.

34
The regional limitation to Peninsular Spanish was a necessary step, despite the current need for similar studies in almost all varieties of Spanish.\textsuperscript{29} Such delimitation provides a more stable basis for diachronic analysis, especially considering the fact that the expansion of the PF seems to be moving at different rates in America than in the Old World (Cartagena 1995-1996). Despite these safeguards against difficulties regarding varietal divergence or regional differences, it is of course acknowledged that the Peninsula itself represents a vast number of Spanish varieties. Regional variation in Peninsular Spanish future expression will be left for future scholars to address.

\subsection*{2.2 Data extraction and exclusions}

The data used for this study include all non-past-tense tokens of the PF and SF found in the corpora studied.\textsuperscript{30} All examples were extracted manually using a careful reading of early texts and computerized searches in texts that were electronically available and that had standard orthography. This original extraction resulted in 5,571 tokens. The count for each form by data set can be seen in Table 2.1 below.

\begin{itemize}
\item \textsuperscript{29} But see, especially, Almeida and Díaz (1998), a variationist study of 20\textsuperscript{th}-century Spanish future expression in the Canary Islands. See also the quantitative studies of Sedano (1994), who examines factors affecting future expression in Venezuelan Spanish; Villa Crésap (1997) on the development of futures in New Mexican Spanish; and Melis (2006), for a diachronic study of the use of PF in Mexican and Peninsular Spanish.

\item \textsuperscript{30} An example of past tense PF is: \textit{Sí, hombre, ¿no te acuerdas que ya me iba a ir} y dije: "Venga, sí, que antes de irme - ""Yeah, man, don't you remember that I was already going to leave and I said: "Oh, yeah, before I go - "" (COREC, CACON012B, 20s). The SF does not have an equivalent past tense form, though the Conditional (e.g. \textit{cantaría 'I would sing'}) may be considered as such (Coleman 1971).
\end{itemize}
Table 2.1. Absolute and relative frequencies of PF and SF by century, raw and normalized per 10,000 words

<table>
<thead>
<tr>
<th>Century (word count)</th>
<th>PF N</th>
<th>Normalized per 10,000</th>
<th>SF N</th>
<th>Norm</th>
<th>Total</th>
<th>PF-SF Ratio (%PF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old Spanish (-311,000)</td>
<td>17</td>
<td>0.5</td>
<td>1855</td>
<td>59.6</td>
<td>1872</td>
<td>1:109 (&lt;1%)</td>
</tr>
<tr>
<td>17-18th c. (-232,000)</td>
<td>59</td>
<td>2.5</td>
<td>1308</td>
<td>56.3</td>
<td>1367</td>
<td>1:22 (4%)</td>
</tr>
<tr>
<td>19th c. (-71,500)</td>
<td>77</td>
<td>10.8</td>
<td>511</td>
<td>71.4</td>
<td>588</td>
<td>1:7 (13%)</td>
</tr>
<tr>
<td>20th c., written (-79,000)</td>
<td>83</td>
<td>10.5</td>
<td>248</td>
<td>31.4</td>
<td>331</td>
<td>1:3 (25%)</td>
</tr>
<tr>
<td>20th c., spoken (-242,000)</td>
<td>830</td>
<td>34.3</td>
<td>583</td>
<td>24.1</td>
<td>1413</td>
<td>1:0.7 (59%)</td>
</tr>
</tbody>
</table>

Total N = 1066

4505

5571

While the data used for this study included all occurrences of both the SF and the PF, regardless of meaning nuance, because my interest was on the constraints on productive use of PF and SF and how these constraints were related to semantic changes, certain uses that were originally extracted from the corpora were identified as irrelevant for this purpose and thus set aside. These uses included the past construction of the SF in *haber + PP* as in (2.5), truncated utterances, as in (2.6) and fixed expressions (cf. Poplack & Turpin 1999:144).\(^{31}\) The fixed expressions that were excluded were, for the 19\(^{th}\) century, *yo apostaré* as in (2.7), which occurred only in first person singular, and for the 20\(^{th}\) century, the expression *si lo sabré yo* 'don't I know it', in (2.8), and syntactically unintegrated uses of *vamos a ver* 'let's see' (cf. Vega Llamas 2002:13; Montes 1962-1963:208-209) and *verás* 'you'll see' (cf. Villa-Crésap 1997; as in (2.9) and (2.10), respectively (cf. Villa-Crésap 1997:52). Interestingly, no fixed phrases were found with *decir* 'say', despite its high frequency and tendency toward epistemic expression in colloquial Spanish (Travis 2005) and its use in fixed SF expressions, for example *cómo le diré* 'how will I tell you', in written and spoken 20\(^{th}\)-century Mexican Spanish (Vega

\(^{31}\)Nominalized uses, such as *el qué dirán* 'the what-will-they-say', were not extracted or counted, though during the extraction process it was noted that these uses were highly infrequent, likely amounting to about .01% of the entire corpus.
Pero habrá pasado frío con un chal. (COREC, CCCON018D, 20s) 'But she must have been (SF) cold with a shawl'

Pues mire, nos va a llamar el -
- Ponme la fecha del Corte Inglés, Antonio. (COREC, CCCON013A, 20s)

'Well look, the [truncated] is going to call (PF) us –
- Put the date for English Court there for me, Antonio'

«Ya entiendo» dixo Sancho: ¿yo apostaré que auia de dezir rata y no gata (Quixote, Cap. VII[1], fol. 24r, 17c.)
'I understand now, said Sancho: I bet (SF) that I should have said rat, and not cat'

¡Ay, que no, que esta vida es un continuo sobresalto, si lo sabré yo! (Pobres diablos, p. 22, 20w)

---

32 Discourse-marker-like uses (i.e. syntactically unintegrated) with ver 'see' removed from the study included va(s) a ver 'you're going to see (PF)' (N=1), vamos a ver 'we'll see'/let's see') (N=51), verás(í) 'you'll see' (N=29), and veremos 'we'll see' (N=4). Syntactically integrated (i.e. non-discourse-marker) uses of these collocations (e.g. le dices: 'Bueno, vale, pues mira, no vas a ver los dibujos' 'you tell him: "Well, fine, well look, you're not going to see (PF) the pictures"' [COREC, CCCON029E, 20s]) were not excluded from study. This exclusion is not meant to imply that no variation occurs in these uses (like in other "fixed" expressions), but only that their use as discourse-markers may affect their distributional patterns in ways that would diverge from other uses of SF and PF. Note, for instance, from the tokens excluded, the tendency for second-person ver to occur in SF (97%, N=29/30) and first-person plural ver to occur in PF (93%, N=51/55). Though the canonical meaning of ver 'see' is not that of a cognitive process or state, it is often used in this way, just as in English. On the formulaic use of vamos a ver in Colombian Spanish, see Montes (1962-1963:208-209), and on ora verás in New Mexican Spanish, see Villa-Crésap (1997:52). These observations lend support to the interpretation of its use in syntactically unintegrated segments as discourse markers, which Travis (forthcoming) shows to have markedly different distributional patterns than their fully verbal counterparts. I have left the close examination of these uses to subsequent research.
'Oh, no, that this life is a continuous somersault, don't I know (SF) it!'

(2.9) Noooooo, no es así. **Vamos a ver**, vamos a ver. (Matar, p. 32, 20w)
'No, it's not like that. Let's see (PF), let's see.'

(2.10) No, no se trata de eso. **Verá...** (Matar, p. 130, 20w)
'No, it's not about that. You'll see (SF)...'

Also excluded were the two tokens of PF in temporal *cuando* 'when' clauses as in (2.11), which do not permit variation with SF and which are distributionally unique in that they cannot co-occur with temporal adverbials.

(2.11) Y **cuando vas a comprar** tú descubres el asunto. (COREC, CCCON019A, 20s)
'And when you are going to buy (PF) you find out about how it is.'

These were included under fixed phrases and nominalizations for the Table 2.2, which shows the number of tokens discarded in each case.

Table 2.2: **Exclusions from general study. token counts**

<table>
<thead>
<tr>
<th></th>
<th>13 PF</th>
<th>13 SF</th>
<th>15 PF</th>
<th>15 SF</th>
<th>17 PF</th>
<th>17 SF</th>
<th>19 PF</th>
<th>19 SF</th>
<th>20w PF</th>
<th>20w SF</th>
<th>20s PF</th>
<th>20s SF</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed/nom.</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>14</td>
<td>--</td>
<td>--</td>
<td>4</td>
<td>12</td>
<td>48</td>
<td>25</td>
<td>104</td>
<td></td>
</tr>
<tr>
<td><em>Haber</em> + PP</td>
<td>--</td>
<td>2</td>
<td>--</td>
<td>5</td>
<td>--</td>
<td>2</td>
<td>--</td>
<td>--</td>
<td>36</td>
<td>--</td>
<td>--</td>
<td>40</td>
<td>85</td>
</tr>
<tr>
<td>Truncated</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>14</td>
<td>19</td>
</tr>
<tr>
<td><strong>Total N</strong></td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>36</td>
<td>4</td>
<td>12</td>
<td>60</td>
<td>84</td>
<td>222</td>
</tr>
</tbody>
</table>

After the exclusions outlined here, I was left with a total of 5,349 tokens. These were then coded according to certain features, as described in the following section.
2.3 Coding and hypotheses

A critical step in carrying out an analysis of variation is selecting the factors (independent variables) to be examined. This choice must be informed, first, by suggestions in previous literature on the factors that may influence the phenomenon (e.g. temporal adverbials), and second, by the insights found in the grammaticization literature, as discussed above (Section 1.2.1) (e.g. retention of older meanings). Through such analysis, we are able to examine at a level of detail never before explored the constraints on variation within the Spanish future temporal reference sector as it developed since the 13th century.

After excluding the examples mentioned in 2.2, all remaining 5,349 tokens were coded (in Excel) according to various factors that were hypothesized to condition PF-SF variation. These factors included verb class, presence of a temporal adverbial and adverbial specificity, subject animacy, clause type, sentence type, polarity, and lexical type, as outlined in Figure 2.2.33

33 As noted in Section 1.1.2.1, chronological, or temporal, distance is a factor that has been repeatedly mentioned as contributing to the choice between PF and SF in Romance futures (e.g. Antonini 1753; Tlaskal 1978; cf. Melis 2006). I have not, however, coded for temporal distance here. In their quantitative, corpus-based study of the effects of this feature on future variant choice in contemporary Canadian French, Poplack and Turpin (1999:150) found that it does not distinguish PF from SF and in their study of English future variation, Walker and Torres Cacoullos (2005) found no correlation between temporal distance and any of the variants. A further reason for omitting this factor is that "temporal proximity turns out to be highly correlated with another factor, type of adverbial specification, such that contexts coded as proximal happen to co-occur with non-specific adverbs, while those coded as distal tend to co-occur with specific adverbs" (Poplack & Turpin 1999:151), though both specific and nonspecific temporal adverbs can occur in both proximate and distal contexts (Walker and Torres Cacoullos 2005). As we will see in Chapter 5, as
Figure 2.2. Contextual features coded

Verb class
Stative (e.g. estar 'be', poder 'be able to', ser 'be', tener 'have')
Motion (e.g. entrar 'enter', salir 'go out')
Ir ('go')
Psychological (e.g. creer 'believe', pensar 'think, saber 'know')
Perception (e.g. oír 'hear', ver 'see')
Dynamic, other (e.g. buscar 'look for', decir 'say', quitar 'take away')

Temporal adverbial/adverbial specificity
None (e.g. seré bonita fiesta 'it'll be (SF) a nice party'; Pobres diablos, p. 18, 20w)
Present, specific (e.g. Pero estaremos aquí la semana que viene 'But I'll be (SF) here next week'; Matar, p. 165, 20w)
Present, nonspecific (e.g. Pronto lo sabremos. 'Soon we will know'; Matar, p. 151, 20w)

Subject animacy
Animate (e.g. Esta tarde nos ocuparemos de África 'This afternoon we'll worry (SF) about Africa'; Billy, p. 38, 20w)
Inanimate (e.g. En estas latitudes jamás puede predecirse cuándo lloverá 'At these latitudes you can never predict when it will rain (SF)'; Algunos modos, El gran cazador blanco, 20w)

Clause type
Main (e.g. No te preocupes hijo, las vengará! 'Don't worry son, I will get revenge for them!'; Billy, p. 73, 20w)
Subordinate (e.g. estoy seguro de que se vengará 'I'm sure that he will get vengeance (SF)'; Pobres diablos, p. 95, 20w)

Sentence type
Declarative (e.g. No lo harás. 'You won't do (SF) it.'; Matar, p. 66, 20w)
Interrogative (e.g. ¿Qué harán? 'What will they do (SF)?'; Billy, p. 61, 20w)

Polarity
Positive (e.g. Les daré un buen susto 'I'll give (SF) them a good scare'; Billy, p. 72, 20w)
Negative (e.g. pero la viudez es un estado que no me disgustará 'but widowhood is a state that will not displease (SF) me'; Billy, p. 35, 20w)

Lexical type
For example,
Lastimar: Cuidado, pequeña, que vas a lastimarte! 'Careful, little one, you're going to hurt (AF) yourself!' (Billy, p. 19, 20w)
Ser: seremos los únicos que escaparemos a la matanza! 'we will be (SF) the only ones that escape the killing!' (Billy, p. 70, 20w)

in Poplack and Turpin's (1999) findings for French, adverbial specification exerts a strong and significant effect on variant choice in future expression in Peninsular Spanish.
These particular contextual features were identified as potential conditioning factors based on: i) the findings of previous quantitative studies on Romance futures; ii) cross-linguistic tendencies found in literature on non-Romance futures, on futures in general, or on tendencies in language change; and iii) the extrapolation of empirical measures that I posited would operationalize the semantic and psychological notions mentioned as affecting variant choice in qualitative studies.

2.3.1 Verb class

The first factor group considered, verb class, was posited the most likely site in which retention of motion meaning in PF would surface. This is because certain classes of verbs, namely stative and psychological verbs, would only be possible in PF if the PF no longer has allative meaning, since motion and stativity are by definition incompatible (e.g. Traugott 2003:365). In Poplack and Turpin’s (1999) study on Canadian French future expression, these tendencies can be seen in their findings on lexical distribution/restriction of the morphological future (SF), which occurs most often (i.e. 22-77% of the time) with the five verbs vouloir 'want', savoir 'know', être 'be', avoir 'have', and revenir 'return', of which the first four are stative or psychological, and least often or not at all (i.e. 0-4% of the time SF) with the seven verbs partir 'leave', mourir 'die', commencer 'begin', falloir 'be necessary', manger 'eat', monter 'ride', s'asseoir 'sit down', and (as they note, the eponym for SF) chanter 'sing' (Poplack & Turpin 1999:156), six (all except falloir) of which are dynamic verbs.

While Poplack & Turpin (1999:156) note that there was no stativity effect in their data, but rather a lexical effect due to fossilization in certain frequently occurring expressions, which they interpret as a sign that the SF is no longer productive in the
speech of French speakers in Ottawa-Hull (cf. Aaron 2003 for similar lexical effects in
the SF of Mexican Spanish and Durán Urrea & Gradoville 2006 on New Mexican
Spanish). If there are, indeed, (highly frequent) fossilized expressions involving statives,
it is likely that when the SF was productive, it also occurred mostly with stative verbs in
what had not yet become fossilized expressions. In Brazilian Portuguese future
expression, verb class effects were found in the selection of variants, where the SF was
consistently favored with non-motion verbs from the 16th-19th centuries (Poplack &
Malvar forthcoming). If the SF is in fact still productive, then, in Peninsular Spanish, it is
still possible that the class of the verb plays a significant role in variant choice in this
case.

2.3.2 Temporal modification and specificity

The second factor, the presence of a temporal adverbial and the specificity of that
adverbial, was aimed at both i) tracing a possible rise in temporality in the PF, and ii)
characterizing the nature of PF future meaning compared to SF in terms of temporal
reference. Adverbial modification has been found to affect PF-SF variation in French
(Flydal 1943; Emirkanian & Sankoff 1985; Jeanjean 1988; Benveniste 1990; Sundell
1991; Vet 1994:59; Helland 1995; Poplack & Turpin 1999) as well as in Spanish (Sedano
1994; Almeida & Díaz 1998). For this study, specific adverbials were considered to be
any within a specified time frame as in (2.12) or that referred to a specific moment as in
(2.13). The most frequent specific temporal adverbials found in the data were mañana
'tomorrow', hoy 'today', specific days (e.g. el viernes), ahora mismo 'right now', and esta
noche 'tonight'.

(2.12) Dentro de una hora se levantará. (COREC, CACON023A, 20s)
'Within an hour he'll get up (SF).'

(2.13) Este mensaje se autodestruirá en diez segundos. (COREC, CPCON006A, 20s)
'This message will self-destruct (SF) in ten seconds.'

Non-specific adverbials were those such as ya 'already', siempre 'always', and un día 'one day', as in (2.14). The most frequently occurring temporal adverbials coded as non-specific were ya 'already/soon', algún día/un día 'someday', luego 'later', después 'later/afterwards', siempre 'always'.

(2.14) Y un día hablaremos del libro de los timos. (COREC, CACON007B, 20s)
'And one day we'll talk (SF) about the book of tricks.'

2.3.3 Grammatical person and animacy

The third factor, grammatical person and animacy, was chosen to operationalize the notion of intention that Bybee and Pagliuca (1987), among others, posit as the intermediate step between obligation and futurity, in the case of SF, and as a possible intermediate step in the case of go-based futures like the PF. Since intention meaning has been associated with both forms at different moments, the hypothesis is that there will be a decrease in first-person singular in the SF as it loses intention meaning, and that in later centuries first-person singular should favor PF, since it developed later than the SF. This operationalization is suggested by (Mellet 1989:277), Bybee, Perkins and Pagliuca (1994:264) and Sedano (1994). Other scholars have attempted to examine this factor in other variationist studies of Romance futures, with inconclusive or at best mostly uninformative results (e.g. Poplack & Turpin 1999; Poplack & Malvar forthcoming)."
Melis (2006:925) claims that she found no first-person or animacy effect in her study of the diachronic development of the PF in Spanish, thus concluding that intentionality did not play a role in early PF. Sedano (1994), on the other hand, found a strong effect of intentionality on PF-SF variation in Venezuelan Spanish, with intention being highly associated with the PF. In the case of both Sedano (1994) and Melis (2006), however, there were differences in methodology and approach. Sedano (1994:235), acknowledging the difficulty of coding for intention in the face of the absence of overt contextual clues, cites Bauhr (1989:91-92, emphasis in original):

Se produce así la situación, al parecer paradójica, de que las dos formas características de que dispone el español para expresar la intención, el futuro en –ré e ir a + infinitivo (desde luego mucho más frecuentes, en esa función, que el explícito tener la intención de + infinitivo) lo hacen, en general, de una manera poca precisa. Lo que es una realidad psicológica dista mucho, en este caso, de ser una realidad lingüística empíricamente comprobable.

Thus, Sedano (1994:235)—apparently somewhat reluctantly—opts for "la única alternative": native speaker intuitions. Sedano offers (2.15) as an example of the way in

manifested as a strong aversion to [Present] in 2nd p. animate contexts (vosé), is not relevant to [20th c. Brazilian Portugués] speech; it neither constrains variant selection nor shows consistent effects” (emphasis mine). Similarly, Poplack and Turpin (1999:154) summarize their findings on person/animacy of subject as follows: “In consideration of the claims that PF is more subjective, occurring preferentially in the context of 1st person subjects, we examined our data according to grammatical person. No particular association of variant with 1st, 2nd or 3rd p. pronouns, full NPs or other types of subject, singular or plural, was detected (contra Söll 1983; Sundell 1991). We did note, however, a distinct increase in IF [or SF] in the context of the formal pronoun of address vous…. The variable rule analysis shows that in fact, IF [or SF] is highly favored in the presence of the formal subject, while PF is correspondingly disfavored here. We take this as an indication of the formal nature of the inflected variant [SF]” (emphasis mine).
which first-person singular PF is associated with intention, and (2.16) as evidence to support her claim that first-person singular SF is not associated with intention.

(2.15) Bueno, por eso yo...yo voy a seguir estudiando. (C113.011)

'Well, that's why I...I am going to keep studying (PF).'

(2.16) Y un señor le dijo: "Mire, usted me da morocotas, que yo le daré...traeré la leche siempre." Y se lo creyó. (C414.134)

'And a man told him: "Look, you give me morocotas [colonial coin], and I'll give (SF) you...I'll bring the milk still." And he believed him.'

In the case of (2.16), Sedano argues that this does not have an intention reading, first, because it is quoted speech and thus removed from the speaker, and second, because the quoted speaker had no intention of carrying out the action. I would argue, however, that the fact that the quoted speaker was believed to have been expressing intention casts doubt upon this analysis. Unfortunately, Sedano does not offer quantitative evidence in favor of her conclusions. In the case of the PF, she notes, "Aunque no es fácil determinar en el corpus cuál es el número de oraciones con sujeto de primera persona singular en las que el FP [PF] trasmite una modalidad de intención, todo parece indicar que la proporción es muy alta," for which she offers three examples (1994:235). Similarly, in the case of the SF, she claims that "no suele haber asociación entre esa forma verbal y la modalidad de intención," for which she offers two examples (Sedano 1994:235-236).
2.3.4 Clause type

The fourth factor, clause type, was aimed at exploring Bybee's (2001) hypothesis that subordinate clauses are more conservative. This idea is supported by Urdiales (1966), who found a tendency for SF to occur in subordinate clauses in the Spanish of León, but contradicted by Almeida and Díaz's (1998:12) finding that the SF is favored in non-subordinate clauses. For this study, coded as subordinate clauses were relative clauses as in (2.17), object complement clauses as in (2.18), clauses introduced by *que* 'that', or related forms, e.g. *aunque* 'although', *porque* 'because', etc. as in (2.19), and certain clauses introduced by *si* 'if' as in (2.20).

(2.17) Pues mire usted, se inicia del siguiente modo, el primer día consta de una visita a, en Valencia, donde *hará* una demostración del aparato de su invención, capaz, según él dice, de curar el S.I.D.A. calentando la sangre de los pacientes. (COREC, CEC0N004D, 20s)

'Well see, it starts in this way, the first day consists of a visit to, in Valencia, where *he will make* (SF) a demonstration of the apparatus of his invention, capable, according to what he says, of curing AIDS heating up the blood of the patients.'

(2.18) ¿Tú crees que *va a aprobar*? (COREC, CAC0N023A, 20s)

'Do you think *he's going to pass* (PF)?'

(2.19) Sujétamela que *voy a sacar* el queso. (COREC, CCC0N034A, 20s)

'Hold it for me (that) *I'm going to take out* (SF) the cheese.'

(2.20) No, no, si *será* estupendo, "-puñetlas!" (COREC, CCC0N037A, 20s)

'No, no, if *it'll be* (SF) stupendous, "-puñetlas!"'

2.3.5 Sentence type

The fifth factor, sentence type, was chosen in order to operationalize the concept of certainty and related meanings, such as volition and determination, cited in various studies (e.g. Bishop 1973:89; Confais 1995; Almeida & Díaz 1998; Jensen 2002). It was hypothesized that interrogatives would indicate less certainty, while declaratives would
indicate more certainty. Various types of clauses were included under the category 'interrogative', including those that may not be considered interrogatives in the traditional sense, but that contributed the same sense of uncertainty. I considered interrogatives to include: wh-questions, as in (2.21); yes-no questions, as in (2.22); subordinate clauses within wh-questions or yes-no questions as in (2.23); and subordinate clauses that included the interrogative (accented) forms of qué 'what', cuándo 'when', etc., without a question mark, as in (2.24).

(2.21) ¿Qué vas a comprarle? (COREC, CBCON048B, 20s)
'What are you going to buy (PF) for her?'

(2.22) ¿Y éste próximo [fin de semana] te vas a a ir también? (COREC, CACON019B, 20s)
'And this next [weekend] are you going to go (PF) too?'

(2.23) ¿Tú crees que volverá a Marbella, Khashoggi? (COREC, CCCON009A, 20s)
'Do you think that Khashoggi will return (SF) to Marbella?'

(2.24) Vamos a ver qué ocurrirá porque - estos rumores - bueno (COREC, CECON006B, 20s)
'Let's see what will happen (SF) because – these rumors – well'

2.3.6 Polarity

The sixth factor, polarity, may seem to be an odd choice to hypothesize as a conditioning factor in PF-SF variation. Though this factor, to the best of my knowledge, is rarely mentioned in qualitative studies of Romance futures (see, however, Franckel 1984 and Vet 1993), quantitative analyses of French future expression have highlighted its contribution to the selection of the SF by speakers (e.g. Deshaies & Laforge 1981; Emirkanian & Sankoff 1985; Lorenz 1989; Lesage 1991; Sundell 1991; Chevalier 1994; Zimmer 1994; Poplack & Turpin 1999). The "spectacular contribution" (Poplack & Turpin 1999:155) of negative polarity to variant choice in Ottawa-Hull French, though
left unexplained, motivated the inclusion of the factor of polarity in the present study. The occurrences that were coded for negative polarity included those negated with *no* 'not', as well as the negators *mucha* 'never', *ni* 'nor' and *jamás* 'never ever'.

### 2.3.7. Lexical type

The seventh factor, lexical type, though not included in the function-based statistical analyses in Chapter 4 due to statistical interactions with the verb class factor group, is meant to test an alternative hypothesis to that tested through the examination of verb class: it is possible that the PF or SF are favored with certain lexical items, yet not with certain verb classes. As noted earlier, Poplack and Turpin (1999) found lexical (but not verb class) effects on the choice of SF in Canadian French, which they interpret as evidence of the non-productivity of this variant. Villa Crésap (1997) also found lexical effects in New Mexican Spanish futurity, and Aaron (2003) in 20th-century Mexican Spanish, due to SF fossilization. By fossilization, I am referring to the non-productive (invariant) use of the SF in common expressions, such as *ya verás* 'you'll see'. My hypothesis here is that Peninsular Spanish will show verb class effects in the selection of future variants SF and PF if these forms are both productive. On the other hand, if the SF is no longer productive, we may find lexical effects like in Villa Crésap (1997), Poplack and Turpin (1999) and Aaron (2003). At the same time, we may also hypothesize that the PF would demonstrate a propensity to occur with certain lexical types in the beginning of its generalization as a future (see Tottie 1991 on lexical diffusion of grammatical structures).
After the tokens were coded according to the seven features in Figure 2.2, I began
the task of delimiting the variable context in preparation for statistical analyses, discussed
in the next section.

2.4 The variable context

2.4.1 Neutralization of meaning in discourse

Before such a study is possible, we must contend with one of the most difficult issues in
studies on syntactic variation. Can syntactic constructions ever truly be considered
alternate ways of saying the same thing in the same way that a fully realized /s/ in [casas]
'houses' means the same thing as an aspirated /s/ in [casah] 'houses' (see Poplack 1980)?
David Sankoff (1988a:153) hypothesizes that the semantic differences between syntactic
constructions in variation, which can be postulated upon reflection, may be neutralized in
certain contexts, i.e. the full panoply of distinctions will not always be salient: "for
certain identifiable sets of alternations, these distinctions come into play neither in the
intentions of the speaker nor in the interpretation of the interlocutor." Such neutralization
is an essential component of language change (Poplack & Turpin 1999:140). The
challenge in a quantitative study of historical syntax is defining the contexts in which
such neutralization is possible (cf. Wolfram 1997:195; Schwenter & Torres Cacoullos
2005).

In phonological studies, such a determination is relatively simple, consisting
principally of the identification of phonetic variants. However, this is not the case with
syntactic variation. In fact, the elusive nature of the parameters of syntactic variation has
led some researchers to express reservations about including grammatical constructions
undergoing semantic change in variation studies at all (e.g. Lavandera 1978; Romaine
1981; but see Silva-Corvalán 1997; cf. Söll 1983:16). Torres Cacoullos (2001) proposes that in studies of variation among grammaticizing forms it is not necessary to consider these forms as meaning the same thing, but rather to understand the common element in a syntactic variable to be the fact that the constructions in variation have overlapping uses that are diachronically related. In the study of future expression, then, we must determine the contexts in which these constructions have overlapped functionally, as well as the contexts in which they have not (cf. D. Sankoff 1988a:153; Silva-Corvalán 1997:119-120; Wolfram 1997:207).

In the study of future expression in Spanish, we are faced with two pressing questions in defining the envelope of variation. First, we must explore the results of the divergence of each one of these constructions in order to define "future" uses versus other uses. Second, we must discover whether the possible semantic differences (described in section 4.1.2) either i) prevent overlap or competition between the forms, or ii) are potentially neutralized in some contexts, allowing variation (cf. Sedano 1994:225). The first of these issues requires form-based analyses, i.e. the study of all occurrences of each form under question, regardless of meaning. The second of these issues demands careful function-based quantitative analysis of the forms in question within the same or similar semantic domains. Thus, as a necessary preliminary step, all uses of SF and PF (except those noted in Section 2.2) were examined in this study (Chapters 3 and 4), allowing for a fuller picture of the synchronic and diachronic lives of these forms, as we will see, so radically transformed since the 13th century, as well as permitting the identification of possible contexts for neutralization, i.e., the envelope of variation.
2.4.2 Defining the envelope of variation

Since this study takes a diachronic perspective on the variation between grammaticizing forms, I believe that the most appropriate approach to the idea of variable context is that it is made up of the functional space in which the PF and the SF have overlapped diachronically (cf. Torres Cacoullos 2001). This space is by no means homogenous; the SF in particular has been assigned various different meanings, including imperative, general truth, intentionality and promises (Section 1.1.2.1; cf. Matte Bon 2005). This diversity in use has been dealt with in different ways in previous quantitative studies of future expression; some (e.g. Walker & Torres Cacoullos 2005; Poplack & Malvar forthcoming) have excluded some of these uses from statistical analyses, since they did not fall into the functional space under consideration and did not clearly allow for variation. However, regarding the abovementioned uses, I have chosen to follow Cartagena's (1995-1996:80) choice to "establecer la oposición simple futuro temporal y de probabilidad," i.e. temporal and epistemic (cf. Villa Crésap 1997, who also took a broad approach). I will justify at some length my definition of this space in Chapters 3, 4, and 6.

Here, suffice it to say that two functions of PF and SF were identified that do not (yet) overlap in Peninsular Spanish (cf. Cartagena 1995-1996:96 on epistemic SF uses), and were thus excluded from the function-based analysis in Chapter 5, but retained for other analyses. These functions are a non-future, motion use of PF, as in (2.25), and a non-future epistemic use of SF, as in (2.26).

(2.25) los domingos **va** la familia **a hacer** la compra (COREC, CCCON034B, 20s) 'on Sundays the family **goes to do** (PF) shopping'
(2.26) *Hará* un par de años o así. (COREC, CACON006D, 20s)
'It must be (SF) a couple of years ago or so'

This delimitation resulted in the exclusion of 243 tokens from the statistical analysis in Chapter 5 (but not from the form-based analyses in Chapters 3 and 4), shown by data set in Table 2.3.

Table 2.3. Exclusions from function-based study, token counts: epistemic SF and motion/habitual PF

<table>
<thead>
<tr>
<th></th>
<th>13</th>
<th>15</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PF</td>
<td>SF</td>
<td>PF</td>
<td>SF</td>
<td>PF</td>
<td>SF</td>
<td>PF</td>
</tr>
<tr>
<td>Epistemic</td>
<td>--</td>
<td>4</td>
<td>--</td>
<td>16</td>
<td>--</td>
<td>22</td>
<td>--</td>
</tr>
<tr>
<td>Motion/habit.</td>
<td>--</td>
<td>--</td>
<td>3</td>
<td>--</td>
<td>--</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>Total N =</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>16</td>
<td>0</td>
<td>22</td>
<td>2</td>
</tr>
</tbody>
</table>

Further justification for the exclusion of epistemic modal SF contexts from the analysis of variation is that the PF has not yet extended to this use, as shown by speaker repair in (2.27) and (2.28), which are the two modal PF tokens listed for the 20th-century spoken data in Table 2.3.35

(2.27) Esto - esto **va a ser** la eh - nombre - - esto es la dirección, digamos -
- La dirección, ¿no? -
- Su nombre - , ese es su nombre -- (COREC, CEC006A, 20s)

'This – this is going to be (PF) the uh - name - - this is the address, we’ll say –
- The address, no?
- Its name -, that's its name'

(2.28) Hay bodegas de Ribera Duero ¿eh?
- No de la calidad del otro vino.
- Ya, ya.
- Hay mucho vino, sí. ¿eh?
- Coño, pues ¿de qué **va a ser** - de qué será eso? ¿Del terreno, del tiempo? ¿o del - ?
- El terreno, el sol - ¡Quién lo sabe! (COREC, CCCON019A, 20s)

35 These examples of PF were also considered truncated utterances, and thus excluded from the entire analysis, but are included here for purposes of illustration.
There are warehouses in Ribera Duero, huh?
- Not of the quality of the other wine.
- Yeah, yeah.
- There is a lot of wine, yeah. Huh?
- Shit, well what is that going to be (PF) – what will that be (SF) from? From the land, from the weather? Or from the - ?
- The land, the sun – Who knows!

In (2.27), the speaker repairs the epistemic modal PF to a present tense form es 'is', and in (2.28) with the SF será 'will be'. No uses of unrepaired epistemic modal PF were found in these data. The appropriateness of this exclusion for these data is examined in Section 6.1.1.

2.5 Variable rule analysis

Once the envelope of variation was delimited, it was possible to embark upon the study of the variation between PF and SF within future contexts. As mentioned above, one of the goals of this study was to discover the constraints that have affected future variation in Spanish since the 13th century. The identification of the linguistic and social constraints on variation is impossible to achieve through casual observation of data, since, as Bayley (2002:118) points out, it is unlikely that any single contextual factor can account for variation. Even in a case such as future expression in Canadian French, where only one contextual factor is shown to account for the great majority of present-day use of French SF, it is not one of the factors repeatedly referenced in qualitative studies, such as present relevance (e.g. Gagnon 1990) or temporal specificity (e.g. Helland 1997:73); instead, surprisingly, variation in Canadian French future expression is highly constrained by polarity, such that SF occurs almost exclusively in negative polarity (Poplack & Turpin 1999). The discovery of such results further demonstrates the value of multivariate quantitative methods.
Statistical programs such as GoldVarb (Rand & Sankoff 1990), a multivariate analysis program "deliberately designed to handle the kind of data obtained in studies of variation" (Bayley 2002:124), are the most efficient and accurate tool in analyzing the multiple factors conditioning variant choice.\footnote{For a description of how to do a variable rule analysis using GoldVarb, see Guy (1993).} This tool allows researchers to consider simultaneously the statistical effect of numerous linguistic or social factors on a speaker's choice to use a form where there is more than one option. In this case, in my sample of over 5,000 tokens, a variable rule analysis would consider all seven factors I coded; it would then tell me: i) whether a particular factor group had a statistically significant effect on variant choice; ii) how much of an effect it had relative to other factor groups; and iii) the particular linguistic contexts that favored more the choice, for example, of the PF. Each context considered must allow for variation, i.e., both variants must occur in the data sample in that context; also, the factors considered must be independent, i.e., they must not interact with each other.

In discussing the results of these analyses, I will be using a specialized vocabulary, as here defined. Variable rule analysis provides researchers with both the magnitude of effect and the hierarchy of constraints. The magnitude of effect has to do with how strong or weak the effect of any given factor group (e.g. co-occurrence with an adverbial) is in predicting the occurrence of a form, and can be measured by the Range. The hierarchy of constraints tells us the the direction of effect, that is, the ordering of the factors within groups from most to least favorable to the variant form chosen (D. Sankoff 1988a, 1988b) (e.g. the absence of adverbial modification most favors PF, followed by specific temporal adverbs referring to a specific time (such as mañana 'tomorrow'), while
non-specific adverbs (such as *algún día* "someday") disfavor PF and favor SF); in other words, it describes the "grammar" (Poplack and Tagliamonte 2001:92-93), so to speak, underlying variation, and is seen in the *Probability weights*. Since an explanation of how to read the results of variable rule analysis is much facilitated by example, I will postpone fuller explanation until Section 5.3.
3 The Synthetic Future

3.1 Origins and development

The lexical origins of the SF are widely known. The modern SF has its roots in the Latin construction INF + habere 'have' (on this topic, see, for example, Coleman 1971; Bello 1984 [1847]; Gili Gaya 1964; Benveniste 1968; Company 1985-1986; Bybee and Pagliuca 1987; Azevedo 1992:307; Molho 1997; Westmoreland 1997). At the inception of the grammaticization of what became the SF in Spanish, Latin already had a synthetic future (e.g. cantabo 'I will sing'). At the same time, periphrastic constructions made up of an infinitive (INF) and a form of habere 'have', in that particular order, originally with a possessive meaning, came to compete with, and eventually replace, the older synthetic form (e.g. Company 1985-1986; Azevedo 1992:307; Almeida & Diaz 1998; Company & Medina Urrea 1999). In (3.29) we see a constructed example from Hopper and Traugott, which they claim shows prior possessive meaning (1993:42). Note that the construction in (3.29), which shows habere + INF, is not the INF + habere construction that grammaticized into the Spanish SF.

(3.29) Haec habeo cantare. (Hopper & Traugott 1993:42)

'I have these things to sing.'

The shift of infinitives in construction with habere, construction that was not attested before the late Republic (Coleman 1971:215), from full verb to auxiliary to affix took place before the birth of Spanish.37 In this construction, it functioned as an auxiliary

---

37 Coleman (1971:216) offers the following as the earliest known example of this construction: habeo etiam dicere quem... de ponte in Tiberim deiicerit (Cic. S. Rosc. 100). For a fuller discussion on the evolution of this form from Latin, see Coleman (1971), Fleischman (1982), or Pinkster (1987).
by the first century B.C., and the construction could function as a future by the third
century A.D. (Westmoreland 1997:383). In the fifth century, it has been suggested that
*habere* still could appear not fused with the infinitive (e.g. *cantar he*), though its bound
morpheme form (e.g. *cantaré*) is already attested in fourth-century Latin and seventh-
century French (Westmoreland 1997:383).\(^{38}\) The synthetic (i.e. bound morpheme) form is
present in the earliest Spanish vernacular texts, such as the tenth-century *Glosas
Emilianenses*, which has forms such as *jras* 'you will go', *faras* 'you will do', and *tornaras*
'you will turn' (C. Lyons 1978:227). However, up until the seventeenth century (C. Lyons
1978:277; Menéndez Pidal 1980:324), the synthetic SF was in variation with an analytic
construction, with an intervening object pronoun between the infinitive and the *habere*
form (e.g. *cantar lo he*).\(^{39}\)

Despite the fact that these lexical origins remain undisputed, the processes that led
to the development of early INF + *habere* constructions into the synthetic form that exists
today have been the site of some contention. Most functionalist accounts agree (e.g.
Fleischman 1982; Bybee & Pagliuca 1987; Bybee, Perkins & Pagliuca 1994; Sedano
1994) that constructions like (3.29) suffered phonological reduction and semantic
generalization through grammaticization, passing through intermediate phases from
possessive to future meaning. These phases would have been the result of a

\(^{38}\) Lyons (1978:227) argues, however, that the first attestation of the modern-day Romance synthetic form
is from the seventh-century *Fredegar’s Chronicle*: "et ille respondabat ‘non dabo’, Justinianus dicebat
‘daras’.

\(^{39}\) See Company (1985-1986, 2006) on why this construction should not be analyzed as a diachronic
predecessor to the synthetic SF *lo cantaré*. See Section 3.2 for a discussion of this and another construction
demonstrating syntactic variability.
conventionalization of common inferences over time (cf. Bybee, Perkins & Pagliuca 1994:25), until they came to be understood as a part of the INF + habere construction's meaning. Such a process would be manifest in gradual semantic change over time, with intermediate phases not unmasking themselves as such until semantic change is complete, when the form appears in contexts that are not compatible with prior meanings. It is these phases that have proven most elusive to identification and, crucially, empirical study.

Regardless of the clarity of the lexical origins of the SF, use of the SF construction with possessive meaning is not apparent in even the earliest Spanish texts; in fact, I found no such use in my 13th and 15th c. data, presumably because such uses would predate my corpus. In an attempt to exemplify early possessive meaning for SF, Hopper and Traugott (1993:43) offer (3.30), which they suggest acquired its obligatory sense from the gerundive, which itself once expressed obligation.

(3.30) Aedem habuit tuendum.
    house had look.after-GER
     'He had a house to look after.'
     (c. 40 BC, Cicero, Ver. II.1,130; cited in Pinkster 1987:208; cited in Hopper & Traugott 1993:43)

Note, however, that the construction in this example (habere + INF) is not the same construction that developed into the Romance SF (INF + habere), and thus does not offer any direct evidence of early SF semantics. A better early example of the SF source construction is seen in (3.31). This example, however, despite the translation offered (Hopper & Traugott 1993:44), which suggests a possessive reading, is not clearly possessive in meaning, since abstract nouns such as truth are not typically possessed.

(3.31) Et si interrogatus fueris, quomodo dicere habes? Veritatem dicere habeo.

    'And if you were asked, what do you have to/will you say? I will have the truth to say/I will speak the truth.'
Coleman, however, argues that the INF + habere construction derived from the "use of the infinitive to signal Purpose or Destination," which, he explains, is well attested in Plautus (1971:216).

The semantic development of the SF (as well as other futures with similar lexical sources) since its putative possessive lexical beginnings has been treated in numerous studies (e.g. Bello 1984 [1847]; Gili Gaya 1964; Company 1985-1986; Bybee & Pagliuca 1987; Bybee, Pagliuca & Perkins 1991; Company & Medina Urrea 1999). It has commonly been proposed that the first semantic step toward futurity taken by the possessive source construction for SF is obligation (e.g. Gili Gaya 1964:165; Fleischman 1982; Bybee & Pagliuca 1987; Hopper and Traugott 1993:42-43). Examples of use of the SF construction in clearly agent-oriented modality contexts, such as obligation (Bybee, Perkins & Pagliuca 1994:177-179), are nearly as scarce as possessive contexts. Villa-Crésap (1997:55) found no examples of this type in New Mexican data, and Montes (1962-1963:176) offered only one example that may have this function in his study of Colombian Spanish. In my data, I too was able to identify only one example, from the mid-14th-century, in which the SF construction seems to have obligation meaning, seen in (3.32).

(3.32) nunca por vigio nin por folgura dexaredes de fazer tales cosas, porque, aun desque vos murierdes, siempre viva la fama de los vuestros fechos. (Lucanor, Exemplo XVI, p. 117, 14c.)

'never due to vice or sloth should you stop (SF) doing such things, so that, even after you are dead, the fame of your deeds may always live.'
Benveniste (1968:89-90) notes that the type of obligation expressed with the SF construction differs from that expressed, for instance, in English *have to*, which places the agent-oriented modality more internally, in the obligated subject. For example, the expression *She has to finish her homework* carries no particular implication about the source of obligation, be it her own desire or sense of duty, or school policy. On the other hand, Spanish SF, notes Benveniste (1968:89-90; cf. Bybee, Perkins & Pagliuca 1994:261-263), expresses "predestination," in which the source of agent-oriented modality is more external to the subject, offering a sense of what "will be done," resembling the English phrase *be to* (Bybee, Perkins & Pagliuca 1994:262), as in *She is to finish her homework*, in which it is an outside force, such as parental pressure or school policy, and not internal drive, that pushes the subject into action.

An important question that linguists interested in the historical semantics of the SF and other futures (e.g. Bybee & Pagliuca 1987; Bybee, Pagliuca & Perkins 1994) have asked is: how does an agent-oriented modality construction come to express futurity? It is clear that this development is not simply a coincidence. Spanish is not unique in its use of an erstwhile agent-oriented modality construction as a future; in their cross-linguistic study of 76 languages, Bybee, Perkins and Pagliuca (1994:258, 263) found that the obligation-to-future grammaticization path is shared by languages as diverse as Inuit, Basque, Buli, Danish, Slave and Haka, and the predetermination path by Tigre, Yagaria, Yessan-Mayo and Maidu. Bybee and Pagliuca (1987) propose that this semantic change came about through the conventionalization of inferences regarding grammatical subject intention.40

---

40 See Section 4.4 for a discussion regarding "intention" meanings for the PF as well.
The "intention" meaning associated with the SF has been among the most contentious of proposed meanings for grammaticizing future forms. As discussed in Section 2.3, the debate over intentionality in the SF (and in some cases, e.g. Sedano 1994; Melis 2006, for the PF) has not focused on whether this notion is plausible or even correct, but rather on whether it is testable. As Bauhr (1989:91-92) aptly notes, "Lo que es una realidad psicológica dista mucho, en este caso, de ser una realidad lingüística empíricamente comprobable." That is, while it seems apparent that some uses of the construction are predictions and others express intentions, the best way to differentiate these uses is not entirely clear. One of the reasons for this difficulty is that these meanings are not (always) mutually exclusive. For example, in their discussion of Middle English shall, which, like the SF, is derived from an obligation construction, Bybee and Pagliuca (1987:114, emphasis added) state, "shall in first person expressions of intention...has become common, and approaches the sense of prediction." They offer the example in (3.33):

(3.33) And I schal ware alle my wyt to wynne me þeder. (Gawain, line 402)
    'And I shall use all my wit to find my way there.'

Methodologically, scholars have addressed this problem in two different ways. On the one hand, those interested in cross-linguistic paths of semantic change (e.g. Bybee & Pagliuca 1987; Bybee, Pagliuca & Perkins 1991) have discussed illustrative examples in order to place them along a diachronic continuum of semantic shift, drawing on evidence from across languages for clues. If the interest, on the other hand, has been the meaning of a particular construction in one language, and how that meaning changed over time,

---

41 Bybee and Pagliuca define prediction as "the assertion by the speaker that a proposition will be true in time subsequent to the moment of speech" (1987:109-110).
then exhaustively coding tokens for subtle semantic distinctions has been less successful. Poplack and Malvar (forthcoming) explain the need for the identification of overt contextual clues in coding data for meaning distinctions when studying the semantics of Romance future expression:

...subtle semantic or pragmatic distinctions in the message the speaker wishes to convey...are difficult, if not impossible, to operationalize: psychological notions like intention, certainty, probability and doubt cannot be identified in the absence of some overt contextual clue. Distinctions that can be operationalized are those based not on the meaning said to be embodied by a variant (which would be circular), but on supporting contextual indicators of that meaning, where present. Thus for example, notions like agentivity, subjectivity and volition, often associated with selection of future variant, can be captured by factors like grammatical person and animacy.

The contextual factor most commonly associated with intention meaning is the first person singular (Mellet 1989:277; Bybee, Perkins and Pagluica 1994:264; Sedano 1994; Villa 1997:64, 96; Poplack & Turpin 1999; Melis 2006; Poplack & Malvar forthcoming), since we are most capable of asserting our own intentions, but not those of others. The compatibility of first person singular with intention meaning can be seen in (3.34), in which the speaker describes his life plan, i.e. the intentions he has regarding family and raising a son.

(3.34) Et esto hecho, casarme he con una mujer muy fervosa, et de grant linaje et noble; et enpreñarse a de un fijo varón complido de sus miembros; et ponerle he muy buen nombre, et enseñarle he buenas costumbres, et castigarlo he de los castigos de los reyes et de los sabios. (Calila, El sueño del religioso, p. 265, 13c.) 'And this done, I will marry (SF) a very beautiful woman, of great lineage and noble; and she will get pregnant with a male child complete in his members; and I will give (SF) him a good name, and I will teach him (SF) good manners, and I will punish (SF) him with the punishments of kings and wisemen.'

Similarly, the speaker in (3.35) expresses her intention to marry in the form of a promise.

(3.35) Haré lo que mi madre me manda, y me casaré con usted. (El sí, Act III, Scene VIII, 18-19c.)
'I will do (SF) as my mother tells me, and I will marry (SF) you.'

Intention meaning with SF also occurs in modern-day speech, shown in (3.36).

(3.36) Mira, yo la verdad es que no sé si me casaré, pero si me caso, creo que me casaré en Madrid. (COREC, CBCON048A, 20s)

'Look, me the truth is that I don't know if I'll get married, but if I get married, I think that I'll get married (SF) in Madrid.'

It is important to keep in mind that not all uses of first person singular have an intention meaning. Note, for example, the first-person singular use in (3.37). Here, the speaker is making a prediction about the future given the actions of her interlocutor. The speaker certainly does not intend to fall dead; on the contrary, she tries to prevent it by attempting to control her interlocutor.

(3.37) No me lo digas; calla, por Dios, que me caeré muerta. (Celestina, Act 15, 15c.)

'Don't say it to me; be silent, by God, for I will fall (SF) dead.'

The key to intentionality in first person singular, then, is that the subject have (at least limited) control over the realization of the future action. In (3.37), this is not the case, since the interlocutor has control over the outcome; in contrast, in (3.35) the speaker does have a choice whether or not to obey her mother, which allows her to speak of her intent (albeit against her desires) to marry.

While not all first-person uses of future forms have intention meaning, it is much more difficult to have intention meaning with second-person or third-person animate subjects, and impossible with inanimate subjects (unless the subject is anthropomorphized) (Bybee & Pagliuca 1987:114). Take, for instance, (3.38), which like (3.34), (3.35) and (3.36) mentions (among other events) a future marriage, but with second-person and third-person subjects. Here, the meaning is clearly one of prediction, not intention; in fact, it reads much like a fortune being told.
(3.38) Si estos preceptos y estas reglas sigues, Sancho, serán luengos tus días, tu fama será eterna, tus premios colmados, tu felicidad indecible, casarás tus hijos como quisiéres, títulos tendrán ellos y tus nietos, viurías en paz, y beneplácito de las gentes, y en los últimos pasos de la vida te alcanzará el de la muerte en vejez suave y madura. y cerrarán tus ojos las tiernas y delicadas manos de tus terceros neteçuelos. (Quixote, Cap. XLII, fols. 160v-161r, 17c.)

'If these precepts and these rules you follow, Sancho, your days will be (SF) long, your fame will be (SF) eternal, your awards high, your happiness beyond words, you will marry (SF) your children as you wish, they and your grandchildren will have (SF) titles, you will live (SF) in peace and among the good will of people, and in the last steps of life death will reach (SF) you, in smooth and mature old age, and the tender and delicate hands of your great-grandchildren will close (SF) your eyes.'

Such examples, coupled by similar observations of the affinity between intention and first person singular (e.g. Mellet 1989:277; Bybee, Perkins and Pagliuca 1994:264; Sedano 1994; Poplack & Turpin 1999; Melis 2006; Poplack & Malvar forthcoming), based on the real-world fact that our own intentions are simply more accessible to us than those of others, has led me to pay particular attention to first-person singular uses in this chapter and throughout this study.

It is from the intention meaning that Bybee and Pagliuca (1987; cf. Bybee, Paglica & Perkins 1991) hypothesize the meaning of prediction (i.e. "the assertion by the speaker that a proposition will be true in time subsequent to the moment of speech") (Bybee & Pagliuca 1987:109-110)) was able to develop. They explain: "For obligation-derived future markers generally, it is only when prediction has become prominent that the use of the marker with inanimate, non-agent subjects becomes common" (Bybee & Pagliuca 1987:114). This point is vital in diachronic analysis, for it provides a testable hypothesis: if the SF is shifting in semantics from more intention meaning to more prediction meaning during the period under study, we should see a drop in the relative
frequency of occurrences with a first-person singular subject and a rise in occurrences with inanimate subjects.

The last stage proposed by Bybee, Pagliuca and Perkins (1991:27-29) in the grammaticization of futures based on agent-oriented modality uses, like the SF, is that in which the gram develops epistemic and imperative uses. This trajectory would place Spanish SF in the last stage of development, given examples such as (3.57) and (3.68).

The good match between Bybee, Pagliuca and Perkin's (1991) model for future development and the Spanish SF leaves us with many answers, but also with many questions. The stages of semantic change are not always easy to identify, since it is obviously not the case that one day the SF means obligation and the next it means intention; Bybee, Pagliuca & Perkins (1991:28) characterize the stages they identify as part of a "continuum," not discrete points. While in many cases diachronic change results in divergence (evident synchronically as polysemy), it also may manifest itself in more subtle ways, such as in syntactic or semantic distributional constraints on the grammaticizing form (cf. Bybee & Pagliuca 1987:112; Hopper 1991; Poplack & Tagliamonte 2001:207-235). For example, Bybee and Pagliuca (1987:112) explain how a lexical source meaning 'obligation' or 'movement', for example, could affect synchronic distribution of future forms:
These original [lexical] senses gradually weaken, and the marker comes to signal prediction, and to be applicable in propositions with any sort of subject[, not only subjects compatible with the original lexical meaning]. However, we would argue that the original sense of these verbs is not lost entirely, but is rather retained in certain contexts, and hence futures from different sources will have different shades or flavors of meaning. [material in brackets added].

These "shades...of meaning" can at times be identified through qualitative analysis of natural data (e.g. Bybee & Pagliuca 1987; Company & Medina Urrea 1999). Other times, these differences are so subtle that they cannot be identified in case-by-case analysis of individual examples with the naked eye; in fact, sometimes it is not even clear if such differences exist in all contexts (Berschin 1986:302; cf. D. Sankoff 1988a; Poplack & Turpin 1999). In such circumstances, factors affecting variant distribution may be discovered through quantitative analysis (e.g. Poplack & Turpin 1999; Torres Cacoullos 2000, 2001; Poplack & Tagliamonte 2001).

To summarize, the development of SF has been hypothesized to have followed the following diachronic path (Bybee, Pagliuca & Perkins 2001:29):

possession > obligation (predestination) > intention > future > epistemic and imperative

The rest of this chapter will pursue the details of this path, with a particular focus on the ways in which quantitative analysis of large sets of data can provide crucial evidence about the nature of change. Before we begin an examination of the evolution of the functions of SF, however, I would like to pause to look at an important formal change that accompanied the SF in its development.
3.2 Changes in form: Increased fixedness and syntactic variability

3.2.1 Increased fixedness

Through the 17th century, the SF morpheme derived from habere demonstrated variability in its boundedness. While in the present Old Spanish data the habere morpheme always appears as bound to the infinitive when no object pronoun was present, as in (3.39), in the presence of an object pronoun, three different word orders were possible as seen in (3.40)-(3.42).

(3.39) **Bound, no co-occurring object pronoun**
Et yo è, dó *fuire* de ti o dó me apartaré? (Calila, La garza, la culebra y el cangrejo, p. 175, 13c.)
'And me, where will I flee (SF) from you or where will I hide out?'

(3.40) **Bound, preceding pronoun**
"Dueña," dijo el rey, "yo *vos mandare* dar vna de las mías, de las mejores que fueren, e mandarvos-he dar todo lo que ouierdes mester." (Zifar, p. 105, 14c.)
"Ma'am," said the king, "I will order (SF) [that they] give *you* one of mine, of the best that there are, and I will order [that they] give you all that you need."

(3.41) **Bound, bound following pronoun**
yo dixe verdat et *mostravela* por prueva (Calila, La mujer y el siervo, p. 187, 13c.)
'I told the truth and I will show it (SF) with proof'

(3.42) **Analytic construction, with pronoun**
"Dueña," dijo el rey, "yo vos mandare dar vna de las mías, de las mejores que fueren, e *mandarvos-he* dar todo lo que ouierdes mester." (Zifar, p. 105; 14c.)
"Ma'am," said the king, "I will order [that they] give you one of mine, of the best that there are, and I will order (SF) [that they] give *you* all that you need."

What is of particular interest here in terms of syntheticization is the option seen in (3.42), where we find a pronoun (*vos*) separating the SF root from its morpheme. The possibility for an intervening pronoun suggests that the SF's status as a fused form was still under negotiation (cf. Bybee, Pagliuca & Perkins 1991:34). Over time, however, this possibility gradually decreased, then disappeared, as shown in Figure 3.3, which depicts the rate of intervening pronouns with the SF from the 13th-17th centuries.
As Figure 3.3 shows, a total of 16% (171/1058) of temporal SF occurrences were analytic in the 13th century. By the 15th century, it was only 6% (48/770), and in the 17th century—the last time period in which this option occurs in these data—it makes up less than 1% (1/1270). These findings are consistent with those of Menéndez Pidal (1944 [1934]) and C. Lyons (1978:277), who also assert that this variant died out after the 17th century. This process could be understood as evidence that the SF became more fixed during the process of grammaticization, through a gradual change that took place over centuries. This process could also be interpreted, however, not as changes that took place within one construction, but rather as the elimination of an analytic construction that was in variation with the SF, an interpretation favored by evidence presented in Company and Medinba Urrea's quantitative study of this variation (1999).

As mentioned in Section 1.1.1, Company and Medina Urrea (1999) have argued that the disappearance of the analytic SF entailed more than simply the loss of a syntactic option: they argue that the loss of this construction was directly tied to its semantic
specialization as a focalizing element, a meaning which became too narrow to be maintained (see also Company 2006). While I have continued to regard the analytic SF as an option within the SF construction for this study, this is not meant to imply that Company's findings may be disregarded. In the next two sections, I will explore two issues. First, here, I will present quantitative evidence that provides support for Company and Medina Urrea's (1999) claim that there were important functional differences between the analytic SF and the synthetic SF in the Medieval Spanish. Second, in Section 3.2.3, I will argue that the specialized meaning of the analytic cantar lo hé did not die with the construction, but rather that the constraints on the analytic variant were transferred to an incipient and quickly moribund construction characteristic especially of the 17th century, cantarélo, in (3.41) above, a process that has also been noted in a diachronic study of variation in future temporal expression in Brazilian Portuguese (Poplack & Malvar forthcoming).

3.2.2 Syntactic variability I: cantar lo hé

In order to begin to explore possible functional or grammatical differences between the analytic SF (cantar lo hé) and synthetic SF (lo cantaré), I undertook a quantitative comparison of the former, as exemplified in (3.42), and its most common competitor in Old Spanish, exemplified in (3.40). I have limited the discussion in this section to SF with future temporal reference (i.e., epistemic uses have been excluded here), since it is the use of SF as a future that is the focus of studies on posited meaning.

Of the SF tokens with co-occurring object or reflexive pronouns in the 15th-century data set (N=253), preposed pronouns as in lo cantaré were 78% (198/253), the analytic cantar lo hé made up 19% (48/253), and the postposed pronoun construction cantarélo comprised 3% (7/253).
differences between these constructions. While the quantitative analysis included both the 13th and 15th centuries, the 15th century is of particular interest because it is the last data set in this corpus that presents a relative frequency of analytic cantar lo hé above 5%; it is in this end stage in which Company suggests the meaning of this construction became particularly narrow (1985-1986, 2006). This quantitative analysis revealed categorical distributional differences between these two constructions.

In the data sets of Peninsular Old Spanish, the contextual range for the analytic cantar lo hé is indeed much narrower than that of the synthetic SF. Table 3.4 shows the relative frequency of the analytic cantar lo hé among SF tokens with co-occurring pronouns in the 13th and 15th centuries. In the 13th century, the analytic cantar lo hé was limited in this data set to positive polarity contexts, in which it occurred categorically. Other contexts, while not categorical, reveal some restrictions on the occurrence of analytic SF, in line with Company and Medida Urrea's finding that "sólo el futuro analítico está asociado a un cierto tipo de sintaxis, mientras que el futuro sintético es mucho más flexible en su distribución" (1999:76). As shown in Table 3.4, the 13th-century analytic cantar lo hé tends not to occur in interrogatives, as in (3.43), with a relative frequency of 10% (N=3), compared to a relative frequency of 42% (N=168) in declarative sentences.

(3.43) e preguntole: "Dueña, dezirme-hedes por que traedes tantos paños?" "Señora," dixo ella, "yo vos lo dire..." (Zifar, p. 103, 14c.) 'and she asked her: "Ma'am, will you tell me (Analyt.) why you have so many napkins?" "Ma'am," she said, "I will tell you..."

Table 3.4 also reveals a tendency for the 13th-century analytic cantar lo hé not to occur in subordinate clauses, as in (3.44), with a relative frequency of 4% (N=6), compared to a relative frequency of 56% (N=165) in main clauses. Note that even in (3.44), the
temporal future construction used closest to the main verb *decir* 'say/tell' is the synthetic SF (which is also negated), not the analytic *cantar lo hé*.

(3.44) *et yo vos digo que non lo fallarán en el cuerpo suyo et fallarlo an en el arca que tenía el su tesoro.* (Lucanor, Exemplo XIIIIO, p. 109, 14c.)  
'and I tell you that they will not find it on his body and they will find it (Analyt.) in the arca that had his treasure'

In the 15th century, the contextual range of analytic SF is ever more restricted.

Here, the categorical limitation to positive polarity remains, and the analytic *cantar lo hé* becomes more contextually restricted (see Company & Medina Urrea 1999 on the distributional limitations of *cantar lo hé* in other contexts not examined here). In the 15th-century data set, the analytic SF did not occur at all in subordinate clauses or with temporal adverbial modification, that is, it occurred only in positive polarity contexts in main clauses with no temporal adverbial modification, as in (3.45).

(3.45) *Agora nos gozaremos juntas, agora te visitaré. Vernos hemos en mi casa y en la tuya.* (Celestina, Act 17, 15c.)  
'Now we will enjoy (Synth. SF) life together, now I will visit you (Synth. SF). We will see each other (Analyt. SF) in my house and in yours'

Furthermore, just as in the 13th century, the 15th-century analytic *cantar lo hé* is less likely than the synthetic SF to occur in interrogative contexts, as in (3.46) (which also carries agent-oriented modality meanings), with a relative frequency of 10% (2/20), compared to a relative frequency of 20% (46/226) in declarative contexts, as in (3.47). The only two tokens of 15th-century analytic *cantar lo hé* in interrogative contexts occur with the first-person singular.

(3.46) *¿ Pues yré, o tornarme hé? ¡ O dubdosa y dura perplexidad! no sé quál escoja por más sano.* (Celestina, Act 4, 15c.)  
'Well shall I go, or shall I go back (Analyt.)? Oh, doubtful and difficult perplexity! I do not know which to choose for the best.
(3.47) *Ven, hermana, a comer, no hagas agora esse plazer a estos locos porfiados; si no, levantarme he yo de la mesa.* (Celestina, Act 9, 15c.)

'Come, sister, to eat, do not give those crazy stubborn ones the pleasure now; if you do not, I will get up' (Analyst.) from the table.'

Table 3.4. *Relative frequency of analytic cantar lo hē among SF tokens with co-occurring pronouns, 13th and 15th centuries*\(^{43}\)

<table>
<thead>
<tr>
<th>Context</th>
<th>13</th>
<th></th>
<th>15</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td><strong>Clause type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main</td>
<td>165</td>
<td>56%</td>
<td>48</td>
<td>27%</td>
</tr>
<tr>
<td>Subordinate</td>
<td>6</td>
<td>4%</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td><strong>Sentence type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Declarative</td>
<td>168</td>
<td>42%</td>
<td>46</td>
<td>20%</td>
</tr>
<tr>
<td>Interrogative</td>
<td>3</td>
<td>10%</td>
<td>2</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Temporal modification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>18</td>
<td>40%</td>
<td>48</td>
<td>22%</td>
</tr>
<tr>
<td>Absent</td>
<td>153</td>
<td>40%</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td><strong>Polarity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>171</td>
<td>47%</td>
<td>48</td>
<td>21%</td>
</tr>
<tr>
<td>Negative</td>
<td>0</td>
<td>--</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>171</td>
<td>40%</td>
<td>48</td>
<td>19%</td>
</tr>
</tbody>
</table>

* Differences marked in bold are statistically significant at the .05 level.

For 13th century, pair-wise comparisons between the analytic *cantar lo hē* and synthetic SF: clause type, \(p < .0000\), Chi-square = 100.6551; sentence type, \(p = .0008\), Chi-square = 11.17554; polarity, \(p < .0000\), Chi-square = 54.03069. No other significant differences at .05 significance level.

For 15th century: clause type, \(p < .0000\), Chi-square = 21.86667; temporal modification, \(p = .0131\), Chi-square = 6.150836; polarity, \(p = .0354\), Chi-square = 4.427154. No other significant differences at .05 significance level.

The analytic *cantar lo hē*’s categorical limitation in these data to main clauses in positive polarity and nearly categorical limitation to declarative contexts (and without temporal adverbials in the 15th-century data), provide strong support, then, for Company’s argument that the analytic *cantar lo hē* and the synthetic SF do not represent the same variant, and that the former grew ever more contextually restricted before its demis in the 17th century (1985-1986, 2006), since these two constructions demonstrate marked distributional differences.

---

\(^{43}\) Table 3.4 does not include the construction *cantarélo* in the synthetic SF count.
3.2.3 Syntactic variability II: cantarélo

As noted in Section 3.2.1 above, the analytic cantar lo hé represented less than 1% (1/1270) of the tokens by the 17th century, and thus was no longer productive in the 17th-century data set.\textsuperscript{44} The analytic cantar lo hé was not the only other syntactic "option" in the 17th century, however. As we saw in Section 3.2.1, there was yet a third possible option for a haber + INF construction: cantarélo, as in (3.41) above. This construction appears even in the earliest data set, in the 13th century, as in (3.48), and constitutes 1% (6/437) of SF tokens that co-occurred with a pronoun (including lo cantaré, cantar lo hé and cantarélo) in the 13th century. In the 15th century, as in (3.49), the relative frequency of cantarélo among SF with co-occurring pronouns reached 3% (7/253), and in the 17th century, in (3.50), 8% (37/465). The cantarélo construction did not appear in the data after the 17th century.

(3.48) Pues si alguno de nos sabe alguna cosa, digalo et nós mostráremoslo al rey
(Calila, La mujer y el siervo, p. 191, 13c.)
'Well if one of us knows something, say it and we will show you (cantarélo) to the king'

\textsuperscript{44} Note further that the inclusion or exclusion of this low-frequency construction in my multivariate analyses of SF-PF variation in Chapter 5, which span the 17th through the 20th centuries, would have no notable effect on the overall results. Whenever Old Spanish data are presented, however, it is important to recall the presence of both the analytic future (cantar lo hé) and, in the 13th-17th centuries, the SF with a postponed pronoun (cantarélo), in the SF count, a presence which has slightly inflated relative frequencies of SF in main clauses, positive polarity, declarative clauses (in Old Spanish), and clauses with no temporal adverbial.

73
(3.49) que si los pidiere, haréle creer que los ha comido. (Celestina, Act 8, 15c.)
'(that) if he requests them, I will make him (cantarélo) believe that he has eaten them'

(3.50) Hablarele, mas no quiero. (Príncipe, Act II, fol. 14v, line 812, 17c.)
'I will speak to him (cantarélo), but I do not want to'

While it is quite possible, and it in fact has been assumed (e.g. Company & Medina Urrea 1999:66), that this word order is simply an older diachronic ancestor of the synthetic SF _lo cantaré_, equally possible is the hypothesis that, like the analytic SF, this short-lived construction was a different gram.

Table 3.5 presents the relative frequency of _cantarélo_ compared to the synthetic SF _lo cantaré_ in the 17th century—the only time period in this data set in which relative frequency exceeded 5%. Here we see a familiar pattern. Like the analytic SF of the 13th and 15th centuries, 17th-century _cantarélo_ occurs categorically in positive polarity contexts, with a relative frequency of 10% (N=38). Also like the analytic SF of Old Spanish, 17th-century _cantarélo_ tends not to occur in subordinate clauses, as in (3.51), at a relative frequency of less than 1% (N=1), compared to a 13% (N=37) relative frequency in main clauses, as in (3.52); furthermore, and _cantarélo_ tends to occur without temporal adverbial modification, with a relative frequency of 10% (N=36) without adverbials, as in (3.53), and 2% (N=2) with adverbials, as in (3.54).

(3.51) Salid, que a vuestra luz, mis dos estrellas, esconderase la envidiosa luna y gozaré mi bien secreto y solo. (Príncipe, Act I, fol. 11r, line 575, 17c.)
'Come out, for in your light, my two stars, the envious moon will hide herself (cantarélo) and I will enjoy my love in secret and alone.'

(3.52) Padre, hareísme mal criado. (Príncipe, Act I, fol. 13r, line 716, 17c.)
'Father, you will spoil me (cantarélo) [Lit. you will make me poorly raised]'

(3.53) ¡Cortaréle aquella lengua! (Dama boba, Act III, Scene XXVII, line 1109, 17c.)
'I will cut out (cantarélo) that tongue of hers!'
(3.54) Vendraste a querer después. (Príncipe, Act II, fol. 5v, line 257, 17c.)
'He will come (cantarélo) to love you later.'

The only context in which the general distributional tendencies of cantarélo do not match those of Old Spanish cantar lo hé is sentence type: cantarélo has a higher relative frequency in interrogatives (20%, N=5), as in (3.55), than in declaratives (8%, N=33), as in (3.56). This difference is statistically significant when considered on its own (see Table 3.5), but multivariate analysis reveals sentence type to not be a statistically significant factor group in the 17th-century data set in relation to the variation between cantarélo and lo cantaré.45

(3.55) Pues, ¿llevaráme a su casa y tendréame allá también? (Dama bobá, Act I, Scene XIV, line 833, 17c.)
'Well, will you take me (cantarélo) to your house and will you have me (cantarélo) there too?'

(3.56) tomarele el pulso y auisaréte si has de venir a estar conmigo o no. (Quixote, Cap. XXXVI, fol. 141v, 17c.)

---

45 A multivariate analysis that included the factor groups of verb class, clause type, sentence type, and temporal adverbial showed that the factor groups clause type and temporal adverbial had significant effects on cantarélo-lo cantaré variant selection in the 17th century Peninsular Spanish data. Main clauses were shown to favor cantarélo with a Probability weight of .78, and subordinate clauses to disfavor cantarélo with a Probability weight of .11, giving a Range of 67. Lack of temporal adverbial modification was shown to slightly favor cantarélo with a Probability weight of .58, and temporal adverbial modification was shown to disfavor cantarélo with a Probability weight of .20, giving a Range of 38. For this analysis, $p = .0006$, Chi-square/cell = 0.3504, Log likelihood = -111.204. Sentence type and verb class were not found to be statistically significant, and polarity was excluded from analysis because the context was invariable. Marginals on sentence type gave a Probability weight of [.67] favoring cantarélo in interrogative contexts, and of [.49] in declarative contexts. See Section 5.3 for an explanation of these terms and of the use of multivariate analyses as a tool for studying variation.
'I will take (cantarélo) its pulse, and I will let you know (cantarélo) if you should come to be with me or not.'

Table 3.5. Relative frequency of cantarélo among synthetic SF tokens with co-occurring pronouns, 17th century

<table>
<thead>
<tr>
<th>Context</th>
<th>N</th>
<th>%</th>
<th>% of data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clause type</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main</td>
<td>37</td>
<td>13</td>
<td>62</td>
</tr>
<tr>
<td>Subordinate</td>
<td>1</td>
<td>&lt;1</td>
<td>38</td>
</tr>
<tr>
<td><strong>Sentence type</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Declarative</td>
<td>33</td>
<td>8</td>
<td>95</td>
</tr>
<tr>
<td>Interrogative</td>
<td>5</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td><strong>Temporal modification</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>2</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>Absent</td>
<td>36</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td><strong>Polarity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>38</td>
<td>10</td>
<td>84</td>
</tr>
<tr>
<td>Negative</td>
<td>0</td>
<td>--</td>
<td>16</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>38</td>
<td>8</td>
<td>100</td>
</tr>
</tbody>
</table>

Pair-wise comparisons between cantarélo and lo cantaré, all groups significant at .05 level. Clause type, $p < .0000$, Chi-square = 22.25783; sentence type, $p = .0265$, Chi-square = 4.925544; temporal modification, $p = .0190$, Chi-square = 5.498676; polarity, $p = .0045$, Chi-square = 8.084863.

Thus, the 17th-century cantarélo is restricted in much the same way, and in the same contexts, as 15th-century Peninsular Spanish cantar lo hé in these data, occurring in main clauses with positive polarity and without co-occurring temporal adverbials. Given Poplack and Malvar's recent finding that constraints may transfer to new variants as forms enter and exit the variable context (forthcoming), I suggest that the cantarélo construction, which was always present in these data, but did not increase in frequency until the 17th century, essentially took over the functional space of cantar lo hé when the latter construction fell out of use in the 17th century.

---

46 Table 3.5 does not include the analytic construction cantar lo he.
While an extensive investigation of distributional, and thus grammatical, differences between syntactic options that have traditionally been included under the umbrella of "Synthetic Future" is beyond the scope of the present study, the examination presented in Section 3.2 reveals a divergence of SF constructions that includes more than just the analytic cantar lo hé and the synthetic lo cantaré (see Company and Medina Urrea 1999; Company 2006, for studies of the meaning differences between these two constructions), but also, at the very least, cantarélo. For the sake of methodological—though not theoretical—analysis, I have included all of the above constructions in my analysis of the SF, since the nature and extent of grammatical differences in apparently similar constructions has yet to be fully accounted for.

3.3 The functions of SF

3.3.1 Futurity and epistemicity

One observable consequence of grammaticization is that lexical forms get fixed in specific potentially grammatical environments, where they take on separate meaning, which Hopper calls "divergence" (1991:22) and Heine and Reh cal "split" (1984:57-59). As forms generalize and develop new meanings, older meanings of the form do not disappear, but rather co-exist with newer meanings. For example, retention (Bybee & Pagliuca 1987:112) of older meanings in the English will (< OE willan 'want') future is evident in the existence of volitional meaning in some expressions, such as If you will help me, we can finish faster (Bybee, Pagliuca & Perkins 1991:28; see also Bybee & Pagliuca 1987) or I will not go (Bishop 1973:88; cf. Bybee, Perkins & Pagliuca 1994:17).
Synchronously, the result of divergence is often in the form of polysemy (Company 2003:13). 47

Indeed, in the case of the SF, as is to be expected with any grammaticizing form, there has been divergence; that is, it is also used in contexts in which it does not describe a future action (e.g. Azevedo 1992:116). 48 The non-future use that has received by far the most attention in the literature is what I have called the epistemic use (e.g. Bello 1984 [1847]:216; Gili y Gaya 1958; Kitova 1986; Azevedo 1992:116; Baena 1996; Butt and Benjamin 1994; Stage 2002; Vega Llamas 2002). This use marks the speaker's degree of commitment to the truth of the proposition (Jespersen 1992 [1924]:313; J. Lyons 1968:308), shown with SF in (3.57). 49

(3.57) Sí, hombre; ya han puesto - ya han puesto algún especial más de él. Hará un par de años o así [que pusieron ese programa]. (COREC, CACON006D, 20s) 'Yeah, man; they've already put – they've already put on some other special of his. It must be (SF) a couple of years ago or something like that [since they put that program on].'

Here, the verb hacer 'make', which would appear in the Present Indicative hace if the speaker were fully committed to the proposition (i.e. 'it was a couple of years ago'),

47 See, however, Section 3.1 on subtle evidence of retention of older meanings throughout the process of divergence.

48 See also Sections 3.2.2-3.2.3 on divergent constructions within SF with future temporal reference.

49 This definition of epistemic modality is not without contention. Like the concepts of "intention" or "present relevance," "degree of commitment" is not available to the analyst in the absence of co-occurring overt contextual clues. For a discussion on definitions of modality, and a convincing argument that defining epistemic modality based on "speaker commitment" is not sufficient when dealing with natural data, see Narrog (2005:678-680). Narrog (2005:679) proposes that epistemic modality should instead be defined as marking the factual status of a proposition as indeterminate; how such marking is to be recognized, however, is left unclear, and thus offers little improvement.
appears in the SF-marked hará, though the time of reference is clearly present/past. This lack of commitment is further indicated by the co-occurrence of o asi 'or something like that'.

Given the cross-linguistic tendency for future tense markers to take on epistemic modal functions (Fleischman 1982:186, 1989:38, cf. Bybee and Pagliuca 1987), this is not unexpected. Similar epistemic modal uses of forms also used for future tense are found in German (Marschall 1999), Romanian (Ripeanu 1994), English (Brisard 1997; Larreya 2000), Italian (Larreya 2000), French (Larreya 2000; Stage 2002), Polish (Tomaszkiewicz 1988), Dutch, Korean, and Quechua (Bybee & Pagliuca 1987:118-119), among others (see also Bybee, Pagliuca & Perkins 1991 for a cross-linguistic survey of the semantics of future forms, including epistemic uses).

The cross-linguistic frequency with which this particular divergence is found is no accident; futurity and modality are intimately (perhaps inextricably) connected. In fact, Narrog (2005:713) suggests that in some cases future markers may be modal markers passing through a phase of futurity. If we understand future-marking as prediction, following Bybee and colleagues (e.g. 1991, 1994), then futures may even be understood as inevitably epistemic: "The prediction sense of future grams may also be considered an epistemic modality because it also has propositional scope and indicates 'the degree of commitment by the speaker to what he says'" (Bybee, Pagliuca & Perkins 1991:24). It has been simultaneously recognized by other scholars that "it is difficult to find future marking that is purely temporal and does not have an element of prediction, or some other modal notion" (Narrog 2005:712; cf. J. Lyons 1977:816; Palmer 1990:137; Veters & Skibinska 1998:255-257).
The path of development of epistemic uses of temporal SF has been explained in more than one way. In the 19th century, Bello (1984 [1847]:216, emphasis added) noted this use in his grammar of the Spanish language, and described the relationship between futurity and probability as a metaphorical one:

La relación de posterioridad se emplea metafóricamente para la consecuencia lógica, la probabilidad, la conjetura. Las formas "cantaré", "cantaría," o "habré cantado" pierden así su valor temporal en cuanto a la relación de que hablamos: el futuro pasa a presente y el pospretérito a pretérito o copretérito; el antefuturo se convierte en antepresente y el antepospretérito en antecopretérito. Parecerá entonces que hay en el verbo una relación de posterioridad que no cuadra con el sentido de la frase, pero realmente no habrá en ella elemento impropio alguno ni ocioso; habrá sólo una metáfora. El verbo se despojará de aquella fuerza de aseveración que caracteriza a las formas del indicativo y en vez de afirmar una cosa como sabida por nuestra propia experiencia o por testimonios fidedignos, la presentará mediante la imagen del futuro, como una deducción o conjetura nuestra, a la que no prestamos entera confianza.

Exactly 140 years later, Bybee and Pagliuca (1987) propose an argument not far from Bello's. Like Bello, they view the epistemic use of futures as a weakened, semantically bleached use of earlier uses. Furthermore, based on cross-linguistic evidence from Spanish, French, Dutch, Korean, Quechua, and English futures, they characterize the probability meaning as having developed out of lexical sources with obligation meanings, thus placing Spanish epistemic SF in the category of "senses which are original to the construction, or which are weakened versions of the original source meaning" (Bybee & Pagliuca 1987:118; cf. Shepard 1982; Bybee & Pagliuca 1985). In a later study, Bybee, Pagliuca and Perkins (1991:27-29) place future forms with epistemic meanings in the last phase of semantic development in the four-step universal diachronic path they propose (possession > obligation (predestination) > intention > future > epistemic and imperative).

80
Taking these analyses into account, and recalling the dual function of what I have called temporal SF as meaning both intention and prediction (see Section 3.3.2), a hypothesis regarding the development of Spanish epistemic SF emerges. We may follow Bybee, Pagliuca and Perkins' (1991) analysis, and propose that the intention meaning of SF was gradually bleached from the semantics of the construction. Such bleaching would leave prediction (a key component of both futurity and epistemic modality). I would then hypothesize, still following Bybee, Pagliuca and Perkins (1991), that the development of epistemic uses of SF would be a relatively late shift, posterior to SF’s conventionalization as a future. In this case, I have chosen to rely upon overt contextual clues present in the examples examined, most particularly grammatical person and subject animacy and future temporal reference, in order to operationalize these concepts. If such a shift, from intention to future prediction to epistemicity, did indeed occur, then, I hypothesize that these changes would be accompanied by:

i. a drop in first-person uses of SF,

ii. a rise in inanimate subjects in SF, and

iii. a rise in contexts involving prediction but not futurity (i.e. epistemic uses).

3.3.2 Distinguishing meanings

The first meaning that comes to mind when we speak of the SF is obviously its use as a future marker. But what exactly does it mean to be a future marker? With the array of meanings described above in Section 3.3.1, clarification is in order. As noted previously, examples of pure obligation meaning for SF—without temporal readings, with the exception of one, (3.32), were not found in my corpus. Thus, such uses have not
concerned me here. What is of concern is the distinction between "intention" and "prediction" uses of SF, which, as noted, is not straightforward. Because of the practical difficulty encountered in coding large quantities of data for such a meaning (Bauhr 1989:91-92; Sedano 1994), and because grammatical person seems to represent an overt contextual clue regarding (the likelihood of) intention meaning (Mellet 1989:277; Bybee, Perkins and Pagluica 1994:264; Sedano 1994; Villa 1997:64, 96; Poplack & Turpin 1999; Melis 2006; Poplack & Malvar forthcoming), I have made the choice here to include all examples that refer to posterior time, whether as intention or prediction, in the count of "temporal" SF. Intention, while not separated from prediction, will be examined through a close look at first-person occurrences of SF.\footnote{Though this decision is not ideal, considering the analyses of, e.g., Bybee and Pagluica (1987) and Bybee, Pagluica and Perkins (1991), who consider intention and prediction to represent two different diachronic stages, the merging of these for statistical analyses of Romance futures, for the same reasons cited here, is not without precedent in recent variationist approaches (see, e.g., Poplack & Turpin 1999; Poplack & Malvar forthcoming).}

A further question regarding the definition of "future" has been that of distinguishing intention/prediction uses from other, more innovative uses, particularly epistemic uses. Sometimes this distinction is simple. For example, the SF may be used temporally, as in (3.58), which illustrates an intention use that has no overtones of epistemic modality and depicts a co-occurring temporal adverbial mañana 'tomorrow'.

(3.58) Entonces te llamaremos mañana a ver si puedes hacerlo o no. (COREC, CPCON006A, 20s)
'So we'll call (SF) you tomorrow to see if you can do it or not.'

However, as discussed earlier, the temporal SF may also have the possibility of epistemic nuances, i.e., we may interpret some occurrences to also communicate the speaker's lack
of commitment to the truth-value of the proposition. This is the case, for example, in
(3.59), indicated by the co-occurring a lo mejor 'maybe'.

(3.59) -Lo que pasa es que yo no sé si sigue habiendo actuaciones martes, miércoles y
  jueves.
  -Seguramente. Lo que pasa que hoy estaré muy lleno a lo mejor.
  -No.
  -¿No?
  -Podemos ir.
  -algún lado.
  -Sí, sí. (COREC, CACON006D, 20s)

   -The thing is that I don't know if there are still shows Tuesday, Wednesday and
Thursday.
   -Surely. The thing is today it'll be (SF) really full maybe.
   -No.
   -No?
   -We can go.
   -somewhere.
   -Yeah, yeah.

In the same light, Matte Bon (2005:np) provides the following example (3.60), apparently
constructed, in which vendrá 'will come (SF)' can be interpreted as both temporal and
epistemic.

(3.60) - ¿Y Pepe? ¿Cómo es que no ha llegado todavía?
   - Se habrá quedado durmiendo y vendrá más tarde. Siempre hace lo mismo.

   - And Pepe? How is it that he hasn't arrived yet?
   - He must have overslept and he'll (probably) come (SF) later. He always does the
same thing.

The possibility for overlapping meanings or interpretations with modality is common
cross-linguistically, and this phenomenon has been referred to as "merger." In her
discussion of this phenomenon in modals, Narrog (2005:684) explains:

"Merger" means that "it is not necessary to decide which meaning is
intended before the example can be understood; with merger the two
meanings involved are not in certain contexts mutually exclusive" (Coates
1983:17).... According to Coates, both epistemic meaning...and [in the
case of modals.] deontic [or agent-oriented] meaning...are not only
available, but may even be simultaneously intended by the speaker. Among the Modern English modals, should and ought to seem to stand out with respect to their potential for overlap (merger) . . . , but in historical perspective these are not isolated cases.

The reason that such mergers are commonplace historically is simple: it is when more than one reading is possible that the potential for semantic change is ripest. As Company (2003:42) notes, "Los signos ambiguos, esto es, aquellos que no mantienen una relación isomórfica o transparente entre forma y significado, las dos caras del signo lingüístico, son más proclives a cambio que los signos que mantienen una relación biunívoca entre sus dos componentes."

Despite the inextricability of futurity and epistemicity, clearly some uses of the SF no longer are future temporal expression. For the interest of this study, all SF uses with possible temporal readings—with or without epistemic readings—were included as "temporal" SF. Based on the findings in various large-scale studies by Bybee and colleagues (e.g. Bybee & Pagliuca 1987; Bybee, Pagliuca & Perkins 1991; Bybee, Perkins & Pagliuca 1994) that epistemic meanings in future expressions are innovative (a claim further validated by my findings here), this broad inclusion is a conservative move, since the semantic change can only be verified when the old meanings (i.e., in this case, intention/prediction) are no longer a possible interpretation. As Traugott and Dasher (2002:11) note, "without some independent evidence such as changes in morphosyntactic distribution, we have no principled way to demonstrate that a linguistic change has taken place." The rest of this chapter, then, is devoted to an examination of such independent evidence. In Sections 3.4.1 and 3.4.2, I will look at the SF's distributional patterns as a future and as an epistemic marker according to the contextual factors of grammatical
subject, verb class, clause type, sentence type, and polarity. In 3.4.3, I will examine the occurrence of other uses of SF that have been mentioned in the literature.

3.4 Characterizing the SF

3.4.1 Temporal uses of SF

One of the first questions we ask when we begin to study variation is: what does each construction do, and where does it occur? At times, the answers can be exciting, when there has been a change in patterns or when a form occurs in unexpected contexts; however, the answers can also be mundane, showing nothing but internal stability, yet these too constitute a necessary preliminary step in discovering the functions of these constructions. In this section, I will show both stability and change in the distributional patterns of SF when it is used to refer to a moment in the future.

Table 3.6 shows the distribution of grammatical person and animacy in temporal SF by century; for the 20th century, 20w indicates the written corpus, and 20s indicates the spoken corpus (see Appendix). Striking here is the overall consistency: most frequent over 700 years is the first-person singular, most commonly followed by third-person singular human and inanimate subjects, then second-person singular. It is especially interesting to note the similarity between the written and spoken data sets from the 20th century, suggesting that popular plays and natural conversational speech may have similar subject distribution.

51 Note the exception of the 15th century, which shows an elevated second-person singular use of 27% (211/770). The reasons for this could not be determined.
Table 3.6. Distribution of subject in temporal SF by data set

<table>
<thead>
<tr>
<th>Subject</th>
<th>13 (N)</th>
<th>15 (N)</th>
<th>17 (N)</th>
<th>19 (N)</th>
<th>20w (N)</th>
<th>20s (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s</td>
<td>30 (320)</td>
<td>28 (221)</td>
<td>36 (453)</td>
<td>29 (124)</td>
<td>20 (42)</td>
<td>21 (81)</td>
</tr>
<tr>
<td>2s</td>
<td>9 (94)</td>
<td>27 (211)</td>
<td>10 (131)</td>
<td>8 (34)</td>
<td>13 (28)</td>
<td>8 (32)</td>
</tr>
<tr>
<td>3s, human</td>
<td>20 (207)</td>
<td>15 (115)</td>
<td>15 (193)</td>
<td>20 (88)</td>
<td>19 (40)</td>
<td>13 (50)</td>
</tr>
<tr>
<td>3s, inanim.</td>
<td>12 (130)</td>
<td>17 (135)</td>
<td>18 (230)</td>
<td>18 (77)</td>
<td>17 (37)</td>
<td>25 (94)</td>
</tr>
<tr>
<td>animals, s&amp;p</td>
<td>3 (34)</td>
<td>&lt;1 (2)</td>
<td>&lt;1 (4)</td>
<td>7 (3)</td>
<td>1 (3)</td>
<td>-- (0)</td>
</tr>
<tr>
<td>1p</td>
<td>8 (83)</td>
<td>4 (31)</td>
<td>6 (83)</td>
<td>8 (34)</td>
<td>16 (34)</td>
<td>14 (54)</td>
</tr>
<tr>
<td>2p</td>
<td>9 (96)</td>
<td>&lt;1 (4)</td>
<td>5 (59)</td>
<td>6 (28)</td>
<td>6 (12)</td>
<td>3 (12)</td>
</tr>
<tr>
<td>3p, human</td>
<td>7 (77)</td>
<td>4 (35)</td>
<td>5 (63)</td>
<td>6 (27)</td>
<td>6 (13)</td>
<td>12 (44)</td>
</tr>
<tr>
<td>3p, inanim.</td>
<td>2 (17)</td>
<td>2 (16)</td>
<td>4 (54)</td>
<td>4 (17)</td>
<td>1 (3)</td>
<td>3 (12)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Temp. SF N</td>
<td>1058</td>
<td>770</td>
<td>1270</td>
<td>432</td>
<td>212</td>
<td>379</td>
</tr>
</tbody>
</table>

2s: 13-15, p < .0000, x² = 109.9352; 15-17, p < .0000, x² = 100.3055
1p: 13-15, p < .0009, x² = 11.11578; 19-20w, p < .0015, x² = 10.04473
2p: 13-14, p < .0000, x² = 63.06263; 15-17, p < .0000, x² = 27.2694
No other significant differences.

Despite this overall stability, a closer look reveals significant changes in both first-person singular and inanimate subject rates, the two contextual factors related to the subject that were identified as possible sites for the identification of semantic change in Section 3.3.1. Let us look first at the use of inanimate subjects. There are two moments in which we see a significant increase in inanimate subject occurrence, one in Old Spanish and one in the 20th century. The 13th-century rate of singular inanimate subjects, at 12% (130/1058), is significantly lower (p ≤ .0027; χ² = 9.031517) than that of the 15th century, at 17% (135/770). This is also true of the 20th-century written rate of 17% (37/212), a rate that is significantly lower at the .05 level (p < .0391; χ² = 4.25621) than the 20th-century spoken rate of 25% (94/379). This suggests that the SF may have undergone an early shift away from intentionality in Old Spanish, as prediction meaning was conventionalized, and later in the 20th century, as epistemic meaning (which is also incompatible with intention) conventionalized.

The distribution of first-person singular subjects points to a similar conclusion. Recall that I hypothesized that a bleeding of intention meaning would be shown if there
was a decline in first-person singular occurrences of SF. Taking into account the figures in Table 3.6, the difference between the combined (written and spoken corpora) 20th century rate of 21% (123/591) and the 19th-century rate of 29% (124/432) represents a statistically significant difference ($p \leq .0036$; $\chi^2 = 8.486129$) (no other first-person singular or inanimate subject differences between data sets were significant; see below Table 3.6 for significant changes in other subjects). This would indicate that a (further) shift away from intention meaning occurred sometime between the 19th and 20th centuries. Another significant difference ($p \leq .0021$; $\chi^2 = 10.52033$) in first-person singular subject rates is found between the 15th and 17th centuries, where there is a rise from 28% (221/770) to 36% (453/1270), followed by a significant decrease ($p \leq .0082$; $\chi^2 = 6.979659$) to the 19th-century rate of 29% (124/432). It is possible that the elevated rate (36%) in the 17th century indicates a short-lived rise in the use of the SF in intention contexts before the SF began to compete with the PF (see Section 5.3.5.1).

The patterns with verb class, however, are not as revealing: the distribution of verb classes with the temporal SF has remained relatively stable since Old Spanish, as shown in Table 3.7. As we can see, the SF has always appeared about half of the time (47%, 1936/4121), with a range of 41-53%, with dynamic verbs, and about a third of the time (31%, 1264/4121) with stative verbs, ranging from 23-34%. Other verb classes, motion, ir, perception, and psychological—singled out for reasons that will become clear in Chapter 5—are negligible, yet equally stable.

---

52 The rate of SF with stative verbs is similar to Vázquez Rozas and García-Miguel’s (2006) findings for Spanish in general, which showed 25.7% stative verbs (20.2% relational and 5.5% existential).
Table 3.7. Distribution of verb class in temporal SF by data set

<table>
<thead>
<tr>
<th>Verb class</th>
<th>13 (N)</th>
<th>15 (N)</th>
<th>17 (N)</th>
<th>19 (N)</th>
<th>20w (N)</th>
<th>20s (N)</th>
<th>Total (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic*</td>
<td>53 (558)</td>
<td>48 (368)</td>
<td>44 (561)</td>
<td>42 (181)</td>
<td>53 (113)</td>
<td>41 (155)</td>
<td>47 (1936)</td>
</tr>
<tr>
<td>Stative</td>
<td>27 (290)</td>
<td>30 (235)</td>
<td>34 (433)</td>
<td>30 (128)</td>
<td>23 (48)</td>
<td>34 (130)</td>
<td>31 (1264)</td>
</tr>
<tr>
<td>Motion</td>
<td>7 (72)</td>
<td>6 (43)</td>
<td>8 (97)</td>
<td>11 (46)</td>
<td>5 (11)</td>
<td>7 (26)</td>
<td>7 (295)</td>
</tr>
<tr>
<td>Ir</td>
<td>3 (29)</td>
<td>3 (26)</td>
<td>2 (30)</td>
<td>3 (13)</td>
<td>1 (3)</td>
<td>6 (22)</td>
<td>3 (123)</td>
</tr>
<tr>
<td>Perception</td>
<td>4 (41)</td>
<td>6 (47)</td>
<td>7 (85)</td>
<td>6 (28)</td>
<td>8 (16)</td>
<td>6 (25)</td>
<td>6 (242)</td>
</tr>
<tr>
<td>Psychological</td>
<td>6 (68)</td>
<td>7 (51)</td>
<td>5 (64)</td>
<td>8 (36)</td>
<td>10 (21)</td>
<td>5 (21)</td>
<td>6 (261)</td>
</tr>
</tbody>
</table>

Total %

100 100 100 100 100 100 100

Temp. SF N = 1058 770 1270 432 212 379 4121

* The Dynamic category does not include motion verbs, which are listed separately.

Another element of consistency is found in SF’s distribution across clause types: since the 13th century, it has occurred approximately two-thirds to three-fourths of the time (71%, 2944/4121), ranging from 67%-84%, in main clauses. This pattern is shown in Table 3.8.

Table 3.8. Distribution of clause type in temporal SF by data set

<table>
<thead>
<tr>
<th>Clause class</th>
<th>13 (N)</th>
<th>15 (N)</th>
<th>17 (N)</th>
<th>19 (N)</th>
<th>20w (N)</th>
<th>20s (N)</th>
<th>Total (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main</td>
<td>67 (713)</td>
<td>74 (580)</td>
<td>66 (845)</td>
<td>84 (363)</td>
<td>78 (165)</td>
<td>73 (278)</td>
<td>71 (2944)</td>
</tr>
<tr>
<td>Subordinate</td>
<td>33 (345)</td>
<td>26 (190)</td>
<td>34 (425)</td>
<td>16 (69)</td>
<td>22 (47)</td>
<td>27 (101)</td>
<td>29 (1177)</td>
</tr>
</tbody>
</table>

Total %

100 100 100 100 100 100 100

Temp. SF N = 1058 770 1270 432 212 379 4121

In Table 3.9, we see that the SF has also tended to occur in declarative clauses, ranging from 84-93% declarative from the 13th through the 19th century. The 20th-century data reveal a slightly elevated proportion of declarative clauses, at 94% (200/212) in written data and 96% (364/379) in spoken. These 20th-century rates of 95% (564/591) declarative

---

53 Obviously, factors that reveal no change tell us very little about the SF; the data for verb class and clause type in this section are presented for informational purposes only, to demonstrate stability in these contexts. Comparisons with the PF later in this chapter and in Chapter 5 will contribute a fuller characterization of the SF.
clauses are significantly higher \((p < .000, \chi^2 = 16.53968)\) than the average rate of 90% (3197/3542) from the 13th through the 19th centuries.

Table 3.9. *Distribution of sentence type in temporal SF by data set*

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>13</th>
<th>15</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
</tr>
<tr>
<td>Declarative</td>
<td>91 (966)</td>
<td>88 (687)</td>
<td>93 (1182)</td>
<td>84 (362)</td>
<td>94 (200)</td>
<td>96 (364)</td>
</tr>
<tr>
<td>Interrogative</td>
<td>9 (92)</td>
<td>12 (83)</td>
<td>7 (88)</td>
<td>16 (70)</td>
<td>6 (12)</td>
<td>4 (15)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Temp. SF N</td>
<td>1058</td>
<td>770</td>
<td>1270</td>
<td>432</td>
<td>212</td>
<td>379</td>
</tr>
</tbody>
</table>

There is a similar tendency to occur in positive polarity, shown in Table 3.10.

Here, SF appears in contexts with positive polarity 82%-88% of the time in written data; in 20th-century speech, however, positive polarity contexts make up 92% (349/379) of SF occurrences, which is significantly higher \((p \leq .0003, \chi^2 = 13.37611)\) that the rate of 82% (174/212) in the written 20th-century data.

Table 3.10. *Distribution of polarity in temporal SF by data set*

<table>
<thead>
<tr>
<th>Polarity</th>
<th>13</th>
<th>15</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
</tr>
<tr>
<td>Positive</td>
<td>83 (882)</td>
<td>86 (676)</td>
<td>85 (1083)</td>
<td>88 (379)</td>
<td>82 (174)</td>
<td>92 (349)</td>
</tr>
<tr>
<td>Negative</td>
<td>17 (176)</td>
<td>14 (94)</td>
<td>15 (187)</td>
<td>12 (53)</td>
<td>18 (38)</td>
<td>8 (30)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Temp. SF N</td>
<td>1058</td>
<td>770</td>
<td>1270</td>
<td>432</td>
<td>212</td>
<td>379</td>
</tr>
</tbody>
</table>

In this section, then, we have uncovered a relatively stable future SF in terms of its distributional patterns, with the majority of notable changes seen in the 20th century: a decrease in first-person singular subjects, an increase in inanimate subjects, and a decrease in interrogatives and negative polarity. We also saw a prior decrease in inanimate subjects in the 15th century. Chapter 6 will examine most of these changes in the larger context of SF semantics and Spanish futurity. Setting aside the 20th-century, we
see a temporal SF that remained relatively stable from Old Spanish through the 19th century.

These numbers in themselves, however, while revealing internal stability or change, reveal little else. The particular rates of occurrence with certain contextual features may not indicate anything particular about the SF’s function, but rather may be simply a characteristic of future tenses in general, or, perhaps, of general trends that are common in other TMA markers. That is, perhaps affirmative polarity occurs approximately 85% of the time in, for example, Preterit and Present as well. In order to arrive at a clearer characterization of the temporal SF and to identify any unique characteristics, it is useful to compare its distribution to other related forms. It may be compared, for example, to non-future uses of the same form, which may tell us something about futurity, or to other future forms, which may tell us something about the SF. The latter will be explored in detail in Chapter 5; the former is the focus of the next two sections.

3.4.2 **Epistemic SF**

It is well-known that the SF has at least one other, non-future use. According to my findings, epistemic modal use of the Spanish SF has occurred since at least the 14th century. Example (3.61), is from *El Conde Lucanor*, written in 1350, and example (3.62) is from the 17th-century *Príncipe Ynocente*.

(3.61) Tal es Dios et los sus fechos, que señal es que poco lo *conoscerá[n]* los [que] mucho fablan en Él. (Lucanor, Tercera parte, p. 290, 14c.)
'Such is God and his deeds, (that) it's a sign that those who speak of him often must know (SF) very little of him'

(3.62) Pues si de mi estás zeloza, ¿de la luna lo *estarás*? (Príncipe, Act I, fol. 3r, line 130, 17c.)
'Well if you're jealous of me, might you be (SF) of the moon as well?'

The existence of epistemic, non-temporal uses of SF is indisputable. That this use has been around for many centuries is also patent. What has yet to be told is how frequent this use has been diachronically, and the nature of its relationship—if any—to the face of future expression. Is it, as both the Hispanic linguistics literature on the matter and grammaticization theory would suggest, a developing use, perhaps even the most common present-day meaning for the SF?
As Figure 3.4 shows, epistemic uses of SF have accounted for a mere 2% of SF occurrences through the 19th century. In 20th-century written data, the proportion of epistemic uses rose to 10%, and in spoken data it comprised a 24% of all SF occurrences. Even if we assume that the epistemic SF occurred 100-150% more in spoken language (based on the 20th-century findings, where spoken data showed a rate of 24% epistemic, compared to 10% in the written data, a difference of 140%)—and thus a spoken rate of 4-5% from the 15th through the 19th centuries—the epistemic SF has been historically unimpressive. The evidence presented here suggests that, while epistemic meanings were always latent in SF in its prime as a future, this potential lay nearly dormant until the 20th century.\(^{54}\) I would argue that this is evidence in favor of Bybee, Pagliuca and Perkins'\(^{54}\)

---

\(^{54}\) This trend, like the encroachment of the PF into the realm of future temporal expression, appears to have gone further in American Spanish. Sedano (1994:231), for instance, found a rate of 58% (148/249) epistemic SF use in Venezuelan spoken Spanish, and Moreno de Alba (1978:98) reported an uncertainty rate of 73% in Mexican Spanish. Villa-Crésap (1997:58) reported a rate of 70% "uncertainty" contexts for SF in New Mexican spoken Spanish, and Durán Urrea and Gradoville (2006), working with a different
(1991) analysis that places epistemic meanings in the last phase of obligation-derived futures. Furthermore, though the obligation meaning had all but disappeared in these data, the long-term existence of epistemic uses, albeit at very low rates, suggests that this potential meaning may indeed be linked to the source construction; in-depth studies of late Latin would help to confirm this.\(^{55}\) Nonetheless, this semantic potential was not exploited by speakers until the 20\(^{th}\) century. Chapter 6 will explore some of the possible reasons for this sudden rise in epistemic uses of the SF.

It is clear from examples such as (3.61) and (3.62) that some epistemic uses of SF are semantically distinct from temporal SF uses. Are these semantically distinct uses also distinctive in their distribution? While handpicked examples of such use can provide some clues into the nature of the epistemic SF, if we wish to characterize such use, it is of particular interest to quantify its distribution in the contexts in which it tends to occur. Through quantification, we can compare epistemic SF and temporal SF beyond

---

\(^{55}\) On the other hand, Durán Urrea and Gradoville (2006) report that 2.8\% (7/247) of the tokens of PF they examined were used with "modal" meaning, which may suggest that this meaning stems from prediction, not obligation (contra Bybee, Pagliuca & Perkins 1991). Take, for instance, the following example, which they cite as "modal" PF: \textit{El que va a saber es mi hermano en Avendale, cuántos son "Who is going to/might know} (PF) is my brother in Avendale, how many they are". In my opinion, this example could also be read as prediction, as in, "If we ask him, I predict that he will know". These data were extracted from recorded oral interviews, and presumably this is a response to an interview question: the speaker may be suggesting that the interviewer contact the brother mentioned in order to obtain the requested information. At these early stages, and with so few tokens, however, it is too soon to be certain whether the PF has or will ever take on epistemic uses as the SF has.

---

corpus of NM Spanish, reported the highest rate, with 79\% epistemic use in their sample of 76 SF occurrences.
individual occurrences, focusing instead on overall tendencies. Such a comparison can help identify differences in distribution that may offer clues into how the SF lost its temporality, as well as about the particular linguistic contexts that may constrain the epistemic use of the SF. However, before we begin the fruitful task of comparison—taken on in Section 3.5—let us take a brief look at the general distribution of epistemic SF.

Table 3.11 shows the distribution of SF subjects by grammatical person and animacy by century. Here we see a strong tendency to occur with third-person inanimate subjects in most data sets, and a clear tendency not to occur with first-person subjects (cf. Matte Bon 2005 on epistemic first-person uses of SF), categorical in the plural. After inanimate subjects, Third-person singular animate subjects are relatively frequent, followed by 2nd-person singular subjects.

Table 3.11. Distribution of subject in epistemic SF by data set

<table>
<thead>
<tr>
<th>Subject</th>
<th>OldSp</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
</tr>
<tr>
<td>1s</td>
<td>--</td>
<td>--</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2s</td>
<td>15 (3)</td>
<td>18 (4)</td>
<td>9 (4)</td>
<td>7 (3)</td>
<td>50 (12)</td>
<td>14 (33)</td>
</tr>
<tr>
<td>3s, human</td>
<td>30 (6)</td>
<td>18 (4)</td>
<td>26 (11)</td>
<td>21 (5)</td>
<td>23 (28)</td>
<td>24 (54)</td>
</tr>
<tr>
<td>3s, anim.</td>
<td>35 (7)</td>
<td>59 (13)</td>
<td>40 (17)</td>
<td>21 (5)</td>
<td>46 (55)</td>
<td>42 (97)</td>
</tr>
<tr>
<td>animals, s&amp;p</td>
<td>5 (1)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>&lt;1 (1)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>l1p</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2p</td>
<td>5 (1)</td>
<td>--</td>
<td>2 (1)</td>
<td>4 (1)</td>
<td>2 (4)</td>
<td>3 (7)</td>
</tr>
<tr>
<td>3p, human</td>
<td>5 (1)</td>
<td>--</td>
<td>7 (3)</td>
<td>--</td>
<td>10 (12)</td>
<td>7 (16)</td>
</tr>
<tr>
<td>3p, inanim.</td>
<td>5 (1)</td>
<td>4 (1)</td>
<td>9 (4)</td>
<td>--</td>
<td>5 (6)</td>
<td>5 (12)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Epist. SF N</td>
<td>20</td>
<td>22</td>
<td>43</td>
<td>24</td>
<td>120</td>
<td>229</td>
</tr>
</tbody>
</table>

Like subject distribution patterns, the verb class results for epistemic SF further suggest a specialized use of this form. As we see in Table 3.12, epistemic SF uses have been nearly categorically restricted to stative and psychological verbs, at an average of 77% (176/229) with statives and 13% (29/229) with psychological verbs, which together...
account for 90% (205/229) of epistemic SF occurrences. This association is most likely linked to aspectual factors: "cuando los consideramos aislados, sin ningún contexto específico, si no hay datos que contradigan dicha interpretación, los verbos utilizados para describir estados o situaciones tienden a proyectarse en el presente, tanto si se hallan en presente como si están en futuro" (Matte Bon 2005). Only in 20th-century spoken data, with the increase in token frequency (N = 120, compared to 20-43 in other data sets), do we see the beginnings of generalization across historically less favorable verb classes, most notably dynamic verbs.

Table 3.12. Distribution of verb class in epistemic SF by data set

<table>
<thead>
<tr>
<th>Verb class</th>
<th>OldSp % (N)</th>
<th>17 % (N)</th>
<th>19 % (N)</th>
<th>20w % (N)</th>
<th>20s % (N)</th>
<th>Total % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic, non-motion -- -- 2 (1) -- 11 (13) 5 (14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stative</strong></td>
<td>65 (13)</td>
<td>91 (20)</td>
<td>93 (40)</td>
<td>58 (14)</td>
<td>74 (89)</td>
<td>77 (176)</td>
</tr>
<tr>
<td>Motion</td>
<td>10 (2)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>4 (5)</td>
<td>3 (7)</td>
</tr>
<tr>
<td>Ir</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>&lt;1 (1)</td>
<td>&lt;1 (1)</td>
</tr>
<tr>
<td>Perception</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>2 (2)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Psychological</td>
<td>25 (5)</td>
<td>9 (2)</td>
<td>5 (2)</td>
<td>42 (10)</td>
<td>8 (10)</td>
<td>13 (29)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Epist. SF N =</td>
<td>20</td>
<td>22</td>
<td>43</td>
<td>24</td>
<td>120</td>
<td>229</td>
</tr>
</tbody>
</table>

Clause type distribution, in Table 3.13, shows that the epistemic SF occurs more in main clauses, at an average of 80% (182/228). These findings differ from Villa-Crésap's (1997:59) finding that "uncertainty" uses of SF in New Mexican Spanish occurred 65% (51/79) of the time in subordinate clauses (cf. Ultan 1978).

Table 3.13. Distribution of clause type in epistemic SF by data set

<table>
<thead>
<tr>
<th>Clause type</th>
<th>OldSp % (N)</th>
<th>17 % (N)</th>
<th>19 % (N)</th>
<th>20w % (N)</th>
<th>20s % (N)</th>
<th>Total % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main</td>
<td>60 (12)</td>
<td>95 (21)</td>
<td>91 (39)</td>
<td>83 (20)</td>
<td>75 (90)</td>
<td>80 (182)</td>
</tr>
<tr>
<td>Subordinate</td>
<td>40 (8)</td>
<td>5 (1)</td>
<td>9 (4)</td>
<td>7 (4)</td>
<td>25 (29)</td>
<td>20 (46)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Epist. SF N =</td>
<td>20</td>
<td>22</td>
<td>43</td>
<td>24</td>
<td>119</td>
<td>228</td>
</tr>
</tbody>
</table>

*One example, not counted here, is used as discourse marker, and is not syntactically integrated.
We also see, in Table 3.14, a consistent tendency to occur in declarative clauses, at an average of 79% (180/229).

Table 3.14. Distribution of sentence types in epistemic SF by data set

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>OldSp 17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
</tr>
<tr>
<td>Declarative</td>
<td>85 (17)</td>
<td>68 (15)</td>
<td>60 (26)</td>
<td>92 (22)</td>
</tr>
<tr>
<td>Interrogative</td>
<td>15 (3)</td>
<td>32 (7)</td>
<td>40 (17)</td>
<td>8 (2)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Epist. SF N = 20 22 43 24 120

Table 3.15 shows an overwhelming tendency to occur in positive polarity contexts; the epistemic SF has tended not to occur, especially before the 20th century, in negative polarity contexts, which have made up a range of 5%-10% of epistemic SF since Old Spanish.

Table 3.15. Distribution of polarity in epistemic SF by data set

<table>
<thead>
<tr>
<th>Polarity</th>
<th>OldSp</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
</tr>
<tr>
<td>Positive</td>
<td>95 (19)</td>
<td>100 (22)</td>
<td>95 (41)</td>
<td>92 (22)</td>
<td>90 (108)</td>
</tr>
<tr>
<td>Negative</td>
<td>5 (1)</td>
<td>--</td>
<td>5 (2)</td>
<td>8 (2)</td>
<td>10 (12)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Epist. SF N = 20 22 43 24 120

All in all, then, the epistemic SF tends to occur in main, declarative, affirmative clauses. Where we see particularly atypical patterning in epistemic SF is in subject and verb class distribution. There is one particular epistemic use of the SF that merits attention: it is what has been called the "concessive" (Bybee, Perkins & Pagliuca 1994:227; cf. Gili y Gaya 1964:165-166; Pedretti 1999), in which the speaker concedes that a certain fact may be true, but then offers some sort of explicit or implicit contradictory or unexpected analysis, as in (3.63)-(3.64).
(3.63) Esso será de cuerpo, madre, pero no de gentileza, no de estado, no de gracia y discreción, no de linaje, no de presumición con merescimiento, no en virtud, no en habla. (Celestina, Act 6, 15c.)

'That [may] be (SF) in body, mother, but not of gentility, not of state, not of grace and discretion, not of lineage, not of deserved presumption, not virtuous, not in speech'

(3.64) «Poeta bien podrá ser», respondió don Lorenço, «pero grande, ni por pensamiento (Quixote, Cap. XVIII, fol. 66r, 17c.)

"A poet he may (SF) well be," answered Don Lorenzo, "but great, not even in thought"

This use is not very common in the data, but it does appear consistently in all time periods since the 14th-15th centuries. Table 3.16 shows the number of occurrences of concessive SF found per time period, as well as the proportion of all epistemic SF uses that are concessive.

Table 3.16. Proportion of concessive SF within epistemic SF by data set

<table>
<thead>
<tr>
<th>Century</th>
<th>N</th>
<th>% concessive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old Spanish</td>
<td>1/20</td>
<td>5</td>
</tr>
<tr>
<td>17th</td>
<td>1/22</td>
<td>4</td>
</tr>
<tr>
<td>19th</td>
<td>7/43</td>
<td>16</td>
</tr>
<tr>
<td>20th w.</td>
<td>2/24</td>
<td>8</td>
</tr>
<tr>
<td>20th s.</td>
<td>3/120</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>14/229</td>
<td>6</td>
</tr>
</tbody>
</table>

As Table 3.16 shows, concessive epistemic uses have remained a minimal yet stable component of SF.

3.4.3 Other proposed SF functions

Apart from the temporal and epistemic uses discussed above, some scholars have proposed other categories for SF usage. These include, most notably, a "general truth" usage and use as an imperative. I will argue that these "meanings," cannot be adequately distinguished from occurrences of SF in contexts that allow a more prototypical future
reading, given that the former is almost always simply a prediction about a posterior event in a stable world where patterns repeat themselves, and the latter is inherently future-referential. As such, I argue that these two uses fall under futurity, in that they both retain the essential semantics of temporal SF, i.e. reference to a posterior time frame (cf. Villa Crésap 1997, who also approaches futurity in this way).

The first phenomenon under discussion here is the use of the SF to describe a "general truth" (Ultan 1978). Bybee and Pagliuca (1987:119) mention this use as deriving not from the original source of a given future (e.g. desire or obligation), but rather from the characteristic of predictability. This use expresses "eternal truths" (Fleischman 1982) or "predictability" (Coates 1983). An example of this "general truth" use can be seen in (3.65).

(3.65) Ca dizien que quien a sí mismo non es leal menos lo será a otro (Calila, Los gatos y el lobo, p. 345, 13c.)

'For they say that he who is not loyal to himself will be (SF) even less so to another'

The English future-marker will offers the same possibility for general truth expression, as can be seen in the gloss of (3.65). Here we see the SF being used with a nonspecific subject (quien a sí mismo non es leal 'he who is not loyal to himself') with reference to a nonspecific point in time (when he is not loyal to another) posterior to a temporally nonspecific point of reference (when he is not loyal to himself). The posteriority of the clause with SF is evidenced in the fact that his disloyalty to others is somehow contingent upon his previous state of disloyalty to himself. In this way, we have the basic meaning of future: prediction about a situation in a time frame posterior to the reference time. As Bybee and Pagliuca (1987:120) point out, "It is certainly a small step from prediction about future time to generalized prediction, but often the 'timelessness' of such statements
may derive more from the context than from the future marker. Note that...all of the
[general truth] sentences cited [in this paper] have indefinite subjects." The generality of
the "general truth" uses of SF, then, is not due to a generalization of SF semantics, but
rather can be attributed simply to its co-occurrence with a nonspecific subject, the scope
of which includes the entire proposition. Further reason not to attribute the semantics of
this construction to the semantics of the SF construction in particular is that the PF can
also occur in this context, as in (3.66)-(3.67), with the same meaning.

(3.66) ¿Ves, Consuelo, que es un hombre? Un hombre enfermo, va a estar enfermo - y
la que también va a andar un poquito de médicos eres tú. (COREC,
CACION024B, 20s)
'See, Consuelo, that he is a man? A sick man is going to be (PF) sick – and who
else is going to be hanging around doctors a bit is you.'

(3.67) O sea, que el au/ el - la persona que ha sido engañada poco va a conseguir.
porque generalmente, el dinero estas personas se lo gastan inmediatamente.
(COREC, CACION007B, 20s)
'Well, that the- the person who has been conned is going to get (PF) very little,
because generally these people spend the money immediately.'

In (3.66), the subject is an nonspecific hombre enfermo 'sick man', preceded by an
indication that the speaker is going to give evidence that the referent is a man, i.e. he
behaves as men generally do. In (3.67), the generality of the nonspecific subject la
persona que ha sido engañada 'the person who has been conned' is reinforced by the co-
occurrence of the adverbial generalmente 'generally'. Since this shift in propositional
meaning cannot be attributed to the SF, but rather to the nature of the subject, in
conjunction with future meaning, the use of the SF in this context is not—at least at this
point in time in Peninsular Spanish—a separate meaning of SF. Nonetheless, it does seem
to be that the SF has traditionally been associated with general truth uses; the use of PF in
this context appears recent, since it was only found in the 20th-century speech data.
Another commonly mentioned phenomenon that occurs with the SF is its use as an imperative (e.g. Keniston 1937:437-438; Kany 1969 [1963]:195; Bybee & Pagliuca 1987:119; Niño-Murcia 1991; Sedano 1994:229; Villa 1997), as in (3.68).  

(3.68) \( Callarás, \) por Dios, o te echaré dende con el diablo! (Celestina, Act 6, 15c.) 'You will be silent (SF), by God, or I will throw you then with the Devil!'  

Bybee and Pagliuca (1987) identify this use as derived from the prediction meaning of future forms. Keniston (1937:437-438) and Kany (1969 [1963]:195) both affirm that imperative uses of the SF can be found in 16\(^{th}\)-century Spanish texts. Going even further back, Niño-Murcia (1992:708) identifies imperative uses of SF in the earliest known Spanish text, as shown in (3.69), and thus concludes that its status as an innovation is not known: "A primera vista este uso [imperativo] del futuro sintético parece una innovación. Sin embargo, es preciso tener en cuenta que tiene una tradición histórica que se remonta a los primeros tiempos de la lengua castellana y formó parte del repertorio verbal desde el Cid [ca. 1307]."  

(3.69) \( vredes, I yazredes \) una noche (El Cid, v.2635, cited in Niño-Murcia 1992:708) 'you will go (SF), you will stay (SF) there one night'  

In the case of SF, and arguably cross-linguistically, the difference between (2\(^{nd}\) person) future and imperative is pragmatic: it boils down to nothing more than a stronger assertion of the proposition (referring to a future action or state of the second person). This use of futures in imperative contexts is common: in a sample of 76 languages, Bybee, Perkins and Pagliuca (1994:210-211) found 13 in which the same morpheme was

\[56\] A case of the conventionalization of the use of SF as an attenuated imperative, as in _leeráslo_ 'please read it', in Andean Spanish in contact with Quechua is examined by Niño-Murcia (1992:707), who explores both internal and external influences on this innovation (see also Toscano 1953; Albor 1973; Yépez Lasso 1984; Niño-Murcia 1988).
reported to have both imperative and future functions. Within the framework of grammaticization, as Bybee and Pagliuca (1987:119) note, "Because commands necessarily refer to future acts and because commanding is imposing a prediction on a second person, the marker used for future prediction is semantically and pragmatically compatible with making commands." In the same vein, Narrog (2005:711-712) argues, "The pragmatic mechanism by which such a change would occur is obvious. Firmly predicting that the addressee is going to do some action in the future invites inferences that the speaker is obliging the addressee to perform that action." While both "simple" future and imperative involve prediction, the latter is essentially agent-oriented (Bybee, Perkins & Pagliuca 1994:177-179), not unlike one of the proposed prior lexical meanings of the SF itself.

Since imperative meaning itself encompasses an array of semantic subtleties (e.g. Niño-Murcia 1992:708), and the imperative is inherently future (e.g. Bybee & Pagliuca 1987:119), identification of imperative uses of SF can get a bit sticky. This is even more so when we take into account interactional factors; Seco (1982 [1954]:76) points to intonation as key to interpretation of imperative future forms, and Jensen (2002) signals the importance of social context on the interpretation of second-person futures: "si es un amigo quien lo dice, será interpretada no como una orden, sino como una instrucción suave, como una invitación." In many cases, separating imperative from non-imperative uses of the SF would require a subjective assessment of the "firmness" (cf. Narrog 2005:711) of the prediction being made. Consider, for a moment, the following examples. In (3.68) above, the SF callarás 'you will be silent' appears quite strongly asserted, and I would consider it a clear example of imperative use of SF. The use of aprenderás 'you
will learn' in (3.70) is, if anything, even more asserted, since it is given during a beating. However, since psychological verbs are less likely to be understood as commands, this highly assertive SF is less imperative-like than (3.68).

(3.70) ¡Toma, toma y toma! Así aprenderás a respetar la ley. (Billy, p. 23, 20w) 'Take this, take this and take this! This way you will learn (SF) to respect the law.'

The importance of the semantics of the verb can also be seen in (3.71) and (3.72). In both, the SF occurs in the context of a request, a likely context for an imperative. In (3.71), descansarás 'you will rest' could be understood either as pure prediction, or, more likely, as a gently asserted prediction, i.e., with imperative overtones. The same cannot be said of estarás mejor 'you will be better' in (3.72), since, while it does make sense to command somebody to rest, it is unlikely that one would command someone to be better. Thus, two occurrences of SF in the same context, with the same overt contextual clues that show the context of a request, may be interpreted differently.

(3.71) ¿Ves, Laura, lo que haces?... ¡Estás toda trémula, demudada, tan pálida!... Ven aquí, bien mio... Descansarás unos instantes, reclinada tu cabeza contra mi pecho. (Conjuración, Act II, Scene III, 19c.) 'See, Laura, what you do?... You're all trembling, speechless, so pale!... Come here, my love... You will rest (SF) for a few moments, your head leaning against my chest.'

(3.72) Es aprensión tuya, Laura mia; yo nunca estoy triste a tu lado. Ven, yo te lo ruego, ven; aquí estarás mejor... ¿no quieres darme ese gusto?... (Conjuración, Act II, Scene III, 19c.) 'It's your apprehension, my Laura; I am never sad at your side. Come, I beg you, come; here you will be (SF) better... don't you want to give me that pleasure?...'

---

57 This is not entirely due to the semantics of estar 'be'. For instance, one can imagine an instance in which a frustrated mother might say to her restless child, ¡Estarás quieto! 'You will be (SF) still!'
In this study, then, the very few apparently (i.e. in this analyst's interpretation) "imperative" or (or, in the case of first-person plural, "exhortative") tokens, both for SF and PF, were considered along with all other tokens with future temporal reference.\textsuperscript{58}

\textsuperscript{58} A preliminary attempt was made to code for unambiguous imperative meaning. For the SF, this preliminary attempt resulted in 14 tokens for the 15th century, 2 for the 17th century, 3 for the 19th century, 5 for the written 20th-century data, and 1 for the 20th-century spoken data. The 13th-century data were not coded for this factor. Both in light of the low token frequency of what appeared to be clear imperatives, and given the concerns I have outlined, such an attempt did not seem scientifically prudent here. Unfortunately, though this use is mentioned often in the literature, it has, to the best of my knowledge, only been quantified in one study (Niño-Murcia 1992 encountered the same situation, finding no prior quantitative information on imperative use of SF in Spanish), perhaps for these same reasons. Villa-Crésap (1997:59, 68), the only quantitative study of the use of Spanish SF in imperative contexts of which I am aware, reports 2% imperative uses in SF and 6% imperative uses for PF. The only other (indirect) clue found in the literature that may possibly refer to the frequency of use of Romance PF and SF in imperative contexts comes from Poplack and Turpin (1999:61), who state, "we make no claims about any non-future uses of [French SF] and [French PF] (which, with the exception of habituality, are in any event, exceedingly rare in the spoken materials we have investigated)." It is not entirely clear, however, if these "non-future" uses included imperatives, which are not listed among the study's exclusions (Poplack & Turpin 1999:143-145); it appears that Poplack and Turpin may have included these, as I have, within future uses. However, Poplack and Malvar (forthcoming) exclude these uses, considering imperatives to be "false futures," but unfortunately provide no figures. It is possible that this exclusion may be behind the unsystematic results these authors found in grammatical person constraints on future temporal expression in Brazilian Portuguese.
3.5 Differentiating temporal and epistemic SF

In this section I will synthesize the results offered in Sections 3.4.1-3.4.2, to answer more definitively the question of distributional, and thus grammatical, differences between temporal and epistemic uses of the SF.\(^{59}\) We will examine in particular the points in which the difference between their distribution patterns is especially salient: subject, verb class, and clause type. We will also see that there are small yet significant differences between epistemic and temporal SF in terms of sentence type distribution and polarity.

3.5.1 Grammatical person and subject animacy

Perhaps the most striking difference between temporal and epistemic SF patterning is found in grammatical person. As Table 3.17 shows, while speakers have tended to use temporal SF 20%-36% of the time with first-person singular subjects, they do not tend to use epistemic SF in the first person singular (0%-9%).\(^{60}\) This trend is attributable to the semantics of epistemicity, which entails a statement about possibility or probability, i.e. prediction. First-person has been linked to intention meaning (e.g. Section 3.1), and is less compatible (though not completely incompatible) with prediction. In fact, of the 499 tokens of 20\(^{th}\)-century spoken SF examined in this study, only 3 (<1%) were first-person singular epistemic uses, all with the verb tener have; these are shown in (3.73), (3.74) and (3.75).

---

\(^{59}\) See, for example, Company and Medina Urrea (1999) for the argument that distributional differences constitute grammatical differences.

\(^{60}\) The epistemic SF did not occur at all in first-person plural in these data (Table 3.11); therefore, it is not analyzed here.
(3.73) - Pues - pues no estás, porque vuestro teléfono es el único que me sé de memoria.
- ¿Sí? Sí estoy!
- O bueno, tendré en la casa - Tendré la q/ - la de tu casa.
- Estoy de Almudena y entre paréntesis, Simón.
- Sí, es que hay otra en León. Hay otra Almudena por ahí.

' - Well – well you guys aren't [at home], because your phone number is the only one I know by heart.
- Yeah? Yes I am [at home]!
- Oh, well, I must have at home – I must have (SF) the one th/ - the one for your house.
- I'm under Almudena and in parentheses, Simón.
- Yeah, it's that there's another one in León. There's another Almudena around there.'

(3.74) porque aquí hay ¿ves? esta es más blanca. Esta es -
- Sí, florescente.
- No, sí, pero lo mejor que hay lo - la florescente blanca. Joder, cuánta iluminación no tendré en el escaparate, que en la tienda si no enciendo los - los flores no me hace falta.
- No hace falta.
- No hace falta. Lo que pasa es que es otro tipo de luz, ¿estás?
- Sí.
- No es la luz clara, clara, clara - del florescente
- Tira así a colorcillo.
- Es que es así.

'because here there is, see? This one is whiter. This one is –
- Yeah, florescent.
- No, yeah, but the best that there is the – the white florescent. Shit, how much lighting I must have (SF) in the display window, 'cause in the store if I don't turn on the – the florescent ones I don't need to.
- You don't need to.
- I don't need to. The thing is, it's another kind of light, get it?
- Yeah.
- It's not bright, bright, bright light – the florescent.
- It's a little bit colored.
- It's that that's how it is.

(3.75) Pero si hicieras las cosas bien a la primera no habría que regañarte a la segunda – .
Es que lo haces a posta para pincharnos a nosotras.
- Sí, encima -
- Debajo.
- No tendré otra cosa que hacer.
- Desde luego que no. Parece que disfrutes, maja.
'But if you did things well at first then you wouldn't have to be scolded the second
time around. It's that you do it on purpose to bug us.
- Yes, on top -
- Below.
- I must not have (SF) anything else to do.
- Of course not. It seems like you enjoy it, jerk.'

_Tener_ 'have' is relatively common in first-person epistemic SF uses in the 19th century as
well, as in (3.76), making up half (2/4) of these occurrences. The only other verb that
appeared in first-person epistemic SF uses in the 19th century was _ser_ 'be'.

(3.76) DON DIEGO.-- ¿Qué dinero tienes ahí?
SIMÓN.-- _Tendré_ unas cuatro o seis onzas. (El sí, Act II, Scene XI, 18-19 c.)

'DON DIEGO: What Money do you have there?
SIMÓN: I must have (SF) four or six ounces.'

Table 3.17. Proportion of 1st-person singular and inanimate singular subjects in
temporal (T) and epistemic (E) SF by data set

<table>
<thead>
<tr>
<th>Subject</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>E</td>
<td>T</td>
<td>E</td>
</tr>
<tr>
<td>1s</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
</tr>
<tr>
<td>Total SF N =</td>
<td></td>
<td></td>
<td>36 (453)</td>
<td>29 (124)</td>
</tr>
<tr>
<td>3s inanim.</td>
<td>14 (265)</td>
<td>35 (7)</td>
<td>18 (230)</td>
<td>59 (13)</td>
</tr>
<tr>
<td>Total SF N =</td>
<td>1840</td>
<td>20</td>
<td>1270</td>
<td>22</td>
</tr>
</tbody>
</table>

While the tendency for speakers not to use epistemic SF with the first person is
readily understood, a bit more puzzling is the tendency we see in Table 3.17 for speakers
to use epistemic SF with singular inanimate subjects. We may at first postulate that the
episemic SF is more likely to occur with a certain _kind_ of inanimate subject, that is, a
subject that would be most consistent with epistemic contexts. A close examination of the
data revealed eight common kinds of inanimate subject: clause, as in (3.77); count noun,
as in (3.78); mass noun, as in (3.79); body part, as in (3.80); abstract, part of the natural

---

61 In the 19th century, the other two tokens (50%, 2/4) occur with _ser_ 'be'; in the 20th-century written data,
the only token of first-person epistemic SF occurs with _ser_.

106
world, time period, and event. While most of these contexts would not be particularly propitious to either future or non-future epistemic readings, we may propose that body part subjects, due to possible metonymy with the speaker, as in (3.80), may be less likely to occur in contexts with non-future epistemic readings, given the tendency for first-person contexts to disfavor non-future epistemic readings (see, for example, Table 3.17).

(3.77) y no sera acertado dexar que nos halle el sol en la calle; mejor sera que nos salgamos fuera de la ciudad (Quixote, Cap. IX, fol. 31 v, 17c.) 'and it won't be right to allow the sun to find us in the street; it will be (SF) better that we leave the city'

Sancho hijo, guia al palacio de Dulcinea; quíca podría ser que la hallemos despierta. (Quixote, Cap. IX, fol. 31 r, 17c.) 'Sancho, son, lead on to Dulcinea's palace: maybe it could be (SF) that we find her awake.'

(3.78) jamás perderá aquella casa el nombre de Celestina (Celestina, Act 15, 15c.) 'that house will never lose (SF) the name of Celestina'

El papel estará muy bien escrito, pero el señor Don Félix es un grandísimo picarón... (El sí, Act III, Scene V, 18-19c.) 'The paper may be (SF) very well written, but Mr. Don Félix is a really picky guy'

(3.79) Yo digo que la agena luz nunca te harás claro si la propria no tienes. (Celestina, Act 2, 15c.) 'I say that external light will never make (SF) you light [i.e. will never shine on you] if you don't have your own.'

Bah, son infelices como estatuas. Será el calor, este maldito bochorno tropical que nos deshidrata el cerebro (Algunos modos, El gran cazador blanco, 20w) 'Hm, they're unhappy like statues. It must be (SF) the heat, this damned tropical suffocation that dehydrates our brains'

(3.80) Gozará mi lastimado corazón, aquel que nunca recibió momento de placer después que aquella señora conoció. (Celestina, Act 6, 15c.) 'My injured heart will feel joy (SF), that one that never received a moment of pleasure after having met that lady.'

Domingo, te decía antes que me imagino que la cara de César será - plena de satisfacción, ¿verdad? (COREC, CECON006B, 20s) 'Domingo, I was telling you before that I imagine that César's face must be (SF) - full of satisfaction, right?'
Abstract subjects are those that consist of an idea or emotion, such as amor 'love' or
responsabilidad 'responsibility', as in (3.81). Time period subjects included, for instance,
día 'day', hora 'hour', and tiempo 'time', as in (3.82). Subjects counted as events included
muerte 'death' and triunfo 'triumph', as in (3.83). Natural world subjects included el sol
'the sun' and la tierra 'the earth', as in (3.84). These last two only occurred with temporal
SF, not epistemic SF.

(3.81) Alegrarte as con tu madre; descansará tu pena. (Celestina, Act 20, 1c.)
'You will become happy with your mother; your sorrow will rest (SF).'

¿Tan grande será la tiranía de la ignorancia, tan común será ya la superfluidad y
el pedantismo que no se atreven los que lloran en silencio esta general corrupción
a declamar altamente contra ella? (Derrota, n.p., 1p.)
'Could the tyranny of ignorance be (SF) so great, and could superfluity and
pedanticity be (SF) so common that those who cry in silence about this general
corruption do not dare to decry this loudly?'

(3.82) Y tú, madre, perdóname, que otro día se verá en que más nos veamos.
(Celestina, Act 4, 1c.)
'And you, mother, forgive me, another day will come (SF) when we see each
other more.'

Hija, ¿por qué tendrás tan pocas horas el día? (COREC, CCCON022B, 20s)
'Daughter, why might the day have (SF) so few hours?'

(3.83) "Cauallero, que sera deste casamiento? Puedese fazer?" (Zifar, p. 79, 14c.)
"Gentleman, what will be (SF) of this wedding? Can it be done?"

(3.84) Piensa bien en esto que dizes, ca a peligro estamos en este lugar; ca si se la mar
tendiere en este lugar, levornos ha nuestros pollos. (Calila, Los tituy y el
mayordomo del mar. p. 164, 13c.)
'Think hard about what you say, because we are in danger in this place; because if
the sea rises in this place, it will take (SF) our chickens from us.'

We may hypothesize that we would see an imbalance in the distribution of the
types of inanimate subjects between temporal and epistemic uses of SF. However, a
comparison of the types of inanimate subjects that appear in each SF use, shown in Table
3.18, provides scant information: both show remarkably similar distributions. Most frequent for both temporal and epistemic SF are clausal subjects (40% (277/696) and 49% (48/97) respectively), followed by count nouns (16% (113/696) and 23% (22/97) respectively); third most frequent are abstract subjects, which occur 16% (111/696) of the time with temporal SF and 9% (9/97) with epistemic SF. These are then followed by mass nouns (4% and 5%), time periods (4% and 5%), and body parts (4% and 1%). Two low-frequency subject types, events and world elements, at 7% (52/696) and 2% (14/696) respectively in temporal SF, do not appear with epistemic SF in these data.

Table 3.18. *Distribution of kinds of inanimate subjects in temporal and epistemic SF*

<table>
<thead>
<tr>
<th>Kind of inanimate subject</th>
<th>Temporal N</th>
<th>Temporal %</th>
<th>Epistemic N</th>
<th>Epistemic %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clause</td>
<td>277</td>
<td>40</td>
<td>48</td>
<td>49</td>
</tr>
<tr>
<td>Count noun</td>
<td>113</td>
<td>16</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>Abstract</td>
<td>111</td>
<td>16</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Event</td>
<td>52</td>
<td>7</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td>Other</td>
<td>48</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Time period</td>
<td>29</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Mass noun</td>
<td>26</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Body part</td>
<td>26</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>World</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>696</strong></td>
<td><strong>100</strong></td>
<td><strong>97</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

If the explanation for the high proportion of inanimate subjects with epistemic SF cannot be found in the nature of co-occurring inanimates, then perhaps we will find the answer in the nature of epistemicity itself. As we saw in Section 3.3.3, the epistemic use of SF occurs most often with stative verbs (65-93%) over seven centuries (see Table 3.12); this, in turn, may affect the kinds of subjects that tend to be used in the epistemic modal context of conjecture.
First, let us take a look at correlations between subject animacy and verb type in SF, in order to establish general tendencies. Table 3.19 shows that, within SF, animate subjects have always tended to co-occur with dynamic verbs, while inanimate subjects have always tended to co-occur with stative verbs. Given the stability of this pattern, it is useful to synthesize the results seen in Table 3.19, shown in Table 3.20. We see that the percentage of animate and inanimate subjects with dynamic and stative verbs are reversed: 51% of animate subjects occur with dynamic verbs, while 64% of inanimate subjects occur with stative verbs.

Table 3.19. Distribution of verb class with singular animate (A) and singular inanimate (I) subjects in all SF by data set

<table>
<thead>
<tr>
<th>Verb class</th>
<th>13</th>
<th>15</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic</td>
<td>A</td>
<td>I</td>
<td>A</td>
<td>I</td>
<td>A</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>41</td>
<td>52</td>
<td>21</td>
<td>51</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>46</td>
<td>22</td>
<td>71</td>
<td>27</td>
<td>65</td>
</tr>
<tr>
<td>Stative</td>
<td></td>
<td></td>
<td>65</td>
<td>25</td>
<td>66</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>66</td>
<td>22</td>
<td>60</td>
<td>28</td>
</tr>
<tr>
<td>Motion</td>
<td>5</td>
<td>10</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>10</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Ir</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>&lt;1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>&lt;1</td>
<td>3</td>
</tr>
<tr>
<td>Percep.</td>
<td>3</td>
<td>--</td>
<td>7</td>
<td>--</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>--</td>
<td>7</td>
<td>--</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Psych.</td>
<td>8</td>
<td>1</td>
<td>9</td>
<td>--</td>
<td>&lt;1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1</td>
<td>9</td>
<td>--</td>
<td>&lt;1</td>
<td>3</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total N</td>
<td>628</td>
<td>125</td>
<td>556</td>
<td>140</td>
<td>243</td>
<td>268</td>
</tr>
</tbody>
</table>
Table 3.20. Distribution of verb class with singular animate and singular inanimate subjects, all data sets

<table>
<thead>
<tr>
<th>Verb class</th>
<th>Total Animate</th>
<th>Total Inanimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic</td>
<td>N: 1314, %: 51</td>
<td>N: 199, %: 25</td>
</tr>
<tr>
<td>Stative</td>
<td>N: 647, %: 25</td>
<td>N: 508, %: 64</td>
</tr>
<tr>
<td>Motion</td>
<td>N: 150, %: 6</td>
<td>N: 58, %: 7</td>
</tr>
<tr>
<td>Ir</td>
<td>N: 78, %: 3</td>
<td>N: 11, %: 1</td>
</tr>
<tr>
<td>Percep.</td>
<td>N: 143, %: 6</td>
<td>N: 11, %: 1</td>
</tr>
<tr>
<td>Psych.</td>
<td>N: 234, %: 9</td>
<td>N: 6, %: &lt;1</td>
</tr>
<tr>
<td>Total</td>
<td>2566, 100%</td>
<td>793, 100%</td>
</tr>
</tbody>
</table>

It appears, then, that the tendency for epistemic SF to occur with inanimate subjects (Table 3.17) is epiphenomenal, due to its association with stative verbs. We find the reason for this pattern in the simple fact that, when we speak about inanimate subjects, we are most likely to discuss states, not actions (at 64% vs. 25%, respectively, due, no doubt to the inherent inaction of most inanimates. Examples of inanimates as subjects of dynamic verbs, such as (3.85), in which *este mensaje* 'this message' is the subject of the dynamic verb *autodestruirse* 'self-destruct', are less common.

(3.85) Si vienes a comer no - este mensaje no sirve para nada. *Este mensaje se autodestruirá en diez segundos.* (COREC, CPCON006A, 20s)

>'If you come to eat it's not – this message is not worth anything. This message will self-destruct (SF) in ten seconds.'

Epistemic meaning occurs more often than temporal meaning with inanimate subjects (Table 3.17) because inanimate subjects occur overwhelmingly in stative contexts—the context most conducive to epistemic interpretations, as in (3.86) and (3.87).

(3.86) Siempre ha sido difícil librarse del poder de Satanás, pero tratándose de una señora... Ya habrás notado que en todos los pactos de Satanás con los humanos, don Diabo siempre pierde... ¡A causa de unas faldas! ¡Será secuela de la ausencia de diablas en el principio de la era de los ángeles subversivos...! (Pobres diablos, p. 38, 20w)
'It has always been difficult to free oneself from the power of Satan, but dealing with a lady... You must have already noticed that in all of Satan's pacts with humans, Mr. Devil always loses... Because of some skirts! It must be (SF) due to the absence of female devils in the beginning of the era of subversive angels...!'

(3.87) Uno que - que crián con agua del mar. Y entonces está salada la uva.  
- Huy madre! ¿Y por dónde es eso?  
- ¿Y aguanta la parra con el agua de mar?  
- Pues eso es - es un vino típico de Portugal  
- Se conoce.  
- pero no sé cuál *será*. Yo no lo he probado. He oído hablar  
- Pues yo no lo he oído.  
- de él pero no lo he probado.  
- Pues *estaré* por el Sur, ¿eh?  
- Luego ahí en Avila en la Sierra esa de Gredos en Pedrobernardo venden un vino especial - - qué bueno! No lo encuentras. (COREC, CCCON019A, 20s)

'One that – that they plant with sea water. And so the grape is salty.  
- Oh jeez! And where is that around?  
- And does the vine do okay with sea water?  
- Well that is – is a traditional wine from Portugal  
- It is known.  
- but I don't know which one *it might be* (SF). I haven't tried it. I've heard tell  
- Well I haven't heard [of] it.  
- of it but I haven't tried it.  
- Well it *must be* (SF) in the South, huh?  
- Then there in Avila in those Gredos mountains in Pedrobernando they sell a special wine – how good! You don't find it.'

3.5.2 Clause and sentence type, polarity

While the differences in distribution of subjects and verb classes within temporal and epistemic SF are most noticeable, there are, in fact, other distributional differences between the contexts in which the use SF favors a future temporal reading and those in which it does not. These smaller differences are found in clause type, sentence type, and polarity. While overall tendencies (i.e. to appear in main declarative clauses with positive polarity) are the same, we do, indeed, find that the strength of these tendencies is not equal in both functions.
For example, Table 3.21 shows that, while in both temporal and epistemic contexts the SF tends not to occur in subordinate clauses, we see a slightly higher relative frequency of the epistemic SF in main clauses. This table summarizes the relative frequency of temporal and epistemic SF in main and subordinate clauses since the 13th century.

<table>
<thead>
<tr>
<th>Clause type</th>
<th>Temporal</th>
<th></th>
<th>Epistemic</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Main</td>
<td>2931</td>
<td>94</td>
<td>181</td>
<td>6</td>
<td>3112</td>
<td>100</td>
</tr>
<tr>
<td>Subordinate</td>
<td>1202</td>
<td>96</td>
<td>47</td>
<td>4</td>
<td>1249</td>
<td>100</td>
</tr>
</tbody>
</table>

$p ≤ .0059, \chi^2 = 7.583052$

Here we see a small (6% vs. 4%) yet statistically significant ($p ≤ .0059, \chi^2 = 7.583052$) difference in the relative frequency of epistemic SF in main vs. subordinate clauses. The reason for this may be related to semantic constraints: the epistemic SF involves the unknown, i.e. a weaker assertion, and thus presumably would be less compatible with certainty. Let us take a closer look at some examples of epistemic SF in subordinate clauses. One epistemic context in which SF tends to occur in the written texts is in si-clauses (i.e. if-clauses) that have no apodosis (19th c.: 25% (1/4); 20w: 50% (2/4)), as in (3.88), from the 19th century, and (3.89), from the 20th century.

(3.88) Me parece que oigo a lo lejos como ruido de remos... ¿Si será ilusión?... No, no hay duda; los latidos de mi corazón me anuncian ya mi dicha y el temblor se apodera de todos mis miembros... (Conjuración, Act II, Scene II, 19c.)

---

62 Though not all data sets show this tendency, it is the Old Spanish data sets that do not. It is important to note the low number of tokens of epistemic SF in these centuries (N=20); results for these centuries must be taken with a grain of salt.

63 This context did not occur in the 20th-century spoken corpus with contexts propitious to epistemic readings.
'It seems to me that I hear in the distance like the sound of oars... If it is (might it be) (SF) an illusion?... No, there is no doubt; the beating of my heart already announce to me my joy and trembling takes over all of my limbs...'

(3.89) La misma señora que negaba a Lucifer, si será impía, observó cuidadosamente al Moratón, ya descendido de su jerarquía, y le censuró que, siendo purpúrado, llève un reloj de oro. (Pobres diablos, p. 50, 20w)
'The very lady who denied Lucifer, though she may be an imp, carefully observed Maratón, having already descended from his hierarchy, and censured him that he should, being a dignitary, take a gold watch.'

Epistemic SF also tends to occur (see Table 3.22 below, and note 64) in object complement clauses with main verbs that are semantically harmonic with uncertainty, such as suponer 'suppose', as in (3.90), imaginarse 'imagine', as in (3.91), and, most commonly, with no saber 'not know', as in (3.92).

(3.90) Supongo que tendrá usted pruebas, testigos de esa... pertinaz persecución. (Matar, p. 68, 20w)
'I suppose that you have (SF) proof, witnesses to that... pertinent persecution.'

(3.91) Satanasito no atinó a curárselos, que no tenía experiencia ni estudios de ninguna clase. Era, lo decía con orgullo, un autodidacta, que, me imagino, será cosa mala usándola él. (Pobres diablos, p. 39, 20w)
'Little Satan was not able to cure them, since he had no experience or schooling of any kind. He was, he said it with pride, a self-taught man, which, I imagine, must be (SF) a bad thing, coming from him.'

(3.92) Yo no lo he visto, ¿eh? No sé cómo será. (COREC, CCCON021B, 20s)
'I haven't seen it, okay? I don't know what it might be (SF) like.'

In the 20th-century spoken corpus, the SF occurs in an object complement clause under 15 different main verbs, shown in Table 3.22.64 Of these: 33% (5/15) express uncertainty,

---

64 Results for earlier data sets were not included in this table due to the low token frequency of epistemic SF in object complement clauses in other data sets. In the 19th century, there were three tokens in this context, with the main verbs pensar 'think', parecer 'seem', and ser cierto 'be true', the last of which is shown in (3.94). In the 20th-century written data, there were two tokens, one with suponer 'suppose' as a main verb, and one with imaginarse 'imagine'. These are shown in (3.90) and (3.91).
imaginarse 'imagine', no saber 'not know', parecer 'seem', preguntarse 'wonder', and suponer 'suppose'; 20% (3/15) express belief, creer 'believe', estar en 'be under the impression', and pensar 'think'; 33% (5/15) express certainty, asegurar 'assure', es que 'it's that', ser claro 'be clear', ser seguro 'be sure' and tener en cuenta 'take into account'.

The verbs decir 'say' and ver 'see' make up the remaining 14% (2/15).

Given what we know of epistemicity, we might hypothesize that epistemic SF would occur relatively less often than temporal SF with verbs indicating certainty.

Noting, however, that the temporal SF has been associated both with certainty and uncertainty, no particular hypothesis can be made regarding temporal SF in the context of uncertainty. In Table 3.22, we see the frequency of each semantic context for temporal and epistemic SF; the number in the table for each verb or expression indicates the number (N) of occurrences of that verb or expression in the respective context. As shown, 88% (15/17) of epistemic SF uses in object complement clauses occur with main verbs that indicate uncertainty. The most common of these is no saber 'not know', accounting for 80% (12/15) of uncertainty uses with epistemic SF. In contrast, temporal SF in object complement clauses occurs in uncertainty contexts 49% (22/45) of the time, also most often with no saber 'not know' (64%, 14/22). Contexts indicating certainty make up 29% (13/22) of temporal SF occurrences in this context, compared to only 6% (1/17) of epistemic SF occurrences.
Table 3.22. Head verbs of object complement clauses with SF, 20th-century speech

<table>
<thead>
<tr>
<th>Head semantics</th>
<th>Temporal SF</th>
<th>Epistemic SF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
</tr>
<tr>
<td>Uncertainty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaginar 'imagine'</td>
<td>49 (22)</td>
<td>88 (15)</td>
<td>60 (37)</td>
</tr>
<tr>
<td>no saber 'not know'</td>
<td>14</td>
<td>12</td>
<td>26</td>
</tr>
<tr>
<td>parecer 'seem'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>preguntarse 'wonder'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>suponer 'suppose'</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Belief</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>creer 'believe'</td>
<td>11 (5)</td>
<td>6 (1)</td>
<td>10 (6)</td>
</tr>
<tr>
<td>estar en 'be under the impression'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>pensar 'think'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Certainty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>asegurar 'assure'</td>
<td>29 (13)</td>
<td>6 (1)</td>
<td>22 (14)</td>
</tr>
<tr>
<td>es que 'it's that'</td>
<td>7</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>ser claro 'be clear'</td>
<td>2</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td>ser seguro 'be sure'</td>
<td>2</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td>tener en cuenta 'take into account'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Perception/Saying</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>decir 'say'</td>
<td>11 (5)</td>
<td>0 (0)</td>
<td>8 (5)</td>
</tr>
<tr>
<td>ver 'see'</td>
<td>3</td>
<td>--</td>
<td>3</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total N =</td>
<td>45</td>
<td>17</td>
<td>62</td>
</tr>
</tbody>
</table>

In fact, the occurrence of epistemic SF in what have traditionally been considered subordinate clauses with main verbs that are not harmonic with uncertainty is exceedingly rare. Only one example of such a use occurred in this sample of 20th-century speech of Peninsular-Spanish, with the expression es que, shown in (3.93), making up 6% (1/17) of the data.65

(3.93) Pero es que de setenta millones de personas - y - no habrá más de cuarenta o cincuenta mil personas haciendo Derecho. (COREC, CCCON013G, 20s)
'But it's that of seventy million people – and – there must not be (SF) more than forty or fifty thousand doing Law.'

---

65 Note, however, that es que clauses, at least in the 20th century, are not truly subordinate, given the discourse uses of this expression.
In the entire corpus, the only other example of epistemic SF used in a subordinate clause with a main verb indicating certainty is in the 19th-century text, *La Conjuración*, with the expression *ser cierto* 'be true', shown in (3.94).

(3.94) Basta pues que el drama histórico posea la condición esencial de reunir la utilidad y el deleite, para que deba hallar en el teatro acogida y aceptación; y cierto que pocas composiciones *habrá* que puedan ser de suyo tan instructivas, y ofrecer al ánimo un desahogo tan apacible. (Conjuración, Apuntes sobre el drama histórico, 19c.)

'It is sufficient then that historical drama possess the essential condition of uniting utility with pleasure, so that it shall find in the theater welcome and acceptance; and it is true that few compositions must exist (SF) that can be as instructive as yours, and offer the soul such a serene rest.'

The restriction of epistemic readings of SF to uncertainty contexts almost exclusively, at least in the case of object complement clauses, may in part explain the higher relative frequency of epistemic SF readings in main versus subordinate clauses (Table 3.21 above).

A further site for differentiation between temporal and epistemic uses of SF is found in their relative occurrence in interrogative clauses. Table 3.23 shows that, while neither temporal nor epistemic SF is more common in interrogative clauses over declarative clauses, a greater proportion of epistemic SF than the temporal SF occurs in interrogative contexts.

Table 3.23. Distribution of sentence type in temporal and epistemic SF by data set

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Old Sp.</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative</td>
<td>90 (1653)</td>
<td>85 (17)</td>
<td>93 (1182)</td>
<td>68 (15)</td>
<td>84 (362)</td>
</tr>
<tr>
<td>Interrogative</td>
<td>10 (175)</td>
<td>15 (3)</td>
<td>7 (88)</td>
<td>32 (7)</td>
<td>16 (70)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total SF N</td>
<td>1828</td>
<td>20</td>
<td>1270</td>
<td>22</td>
<td>432</td>
</tr>
</tbody>
</table>
The association of epistemic SF with interrogatives is more clearly depicted in Table 3.24, which shows the relative frequency of temporal and epistemic SF within interrogative and declarative sentence types for all data sets.

Table 3.24. Relative frequency of temporal and epistemic SF by sentence type, all data sets

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Temporal</th>
<th>Epistemic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interrogative</td>
<td>354 88</td>
<td>49 12</td>
<td>403 100</td>
</tr>
<tr>
<td>Declarative</td>
<td>3781 95</td>
<td>180 5</td>
<td>3961 100</td>
</tr>
</tbody>
</table>

$p < .0000, x^2 = 42.65468$

Here we see that, epistemic uses make up 5% (180/3961) of SF declaratives but 12% (49/403) of interrogatives; this constitutes a significant difference in temporal and epistemic SF distribution ($p < .0000, x^2 = 42.65468$). This tendency may be due to the fact that interrogative contexts are ideal for emphasizing epistemic modality because some interrogatives themselves express this same modality (Narrog 2005:679-680). Here, then, the line between SF semantics and the meaning of the construction in which it occurs becomes less definite. We can see this in (3.95), in which the speaker explains how foreign journalists expressed doubt (recelo) through questioning.

(3.95) Primero, cuando se trasladó al Jarama este gran premio de - de Europa, de pri/ - inicialmente de Yugoslavia, ahora - denominado gran premio de Europa, hubo - y los lo digo, bastante recelo: "Bueno, pero ¿el Jarama estará en condiciones?". Nos preguntaban los periodistas extranjeros, y nosotros dijimos: "Bueno, eh - pu/ - pu/ - pues claro". Pasamos las manos en el fuego por el Jarama, ¿no? (COREC, CECON006B, 20s)

'First, when this big award from - from Europe was transferred, at fi/ - - initially from Yugoslavia, now - called great award of Europe, there was - , I'm telling you guys, a lot of doubt: "Well, but will Jarama be (SF) in good enough condition?" The foreign journalists would ask us, and we uh - said: "Well, uh - we/ - we/ - well of course". We put our hands in the fire for Jarama, no?'

It is in contexts such as this one that the lack of certainty expressed through interrogatives could reinforce the lack of certainty expressed in (epistemic uses of) SF, thus bringing lack of certainty into salience, and backgrounder future temporal meanings. If frequent
enough, occurrence in interrogative contexts could allow for semantic enrichment through the conventionalization of the use of this construction in epistemic contexts. Though interrogatives represent a highly harmonic context for epistemicity, the proportion of interrogatives is not as high as one might expect if this were the context for conventionalization (see Table 3.23). Furthermore, it is also unlikely that this context has played a key role since it makes up such a small percentage of the data, at 9% (403/4364).

The last contextual feature examined, polarity, also shows small yet statistically significant distributional differences between temporal and epistemic SF. While we have seen that both uses occur in their great majority in positive polarity contexts, Table 3.25 reveals that the tendency not to occur in negative polarity contexts is stronger—at times even categorical—in the epistemic SF. Only in 20th-century spoken data do negative polarity contexts occur with both uses of SF at about equal rates. In all other data sets, negative polarity contexts always surpass 10% for temporal SF, yet fail to reach 10% for epistemic SF.

Table 3.25. Distribution of polarity in temporal and epistemic SF by data set

<table>
<thead>
<tr>
<th>Polarity</th>
<th>Old Sp</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T (N)</td>
<td>E</td>
<td>T</td>
<td>E</td>
<td>T</td>
</tr>
<tr>
<td>Positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>84 (1558)</td>
<td>85 (1083)</td>
<td>100 (22)</td>
<td>88 (379)</td>
<td>95 (41)</td>
<td>82 (174)</td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 (270)</td>
<td>15 (187)</td>
<td>--</td>
<td>12 (53)</td>
<td>5 (2)</td>
<td>18 (38)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total SF N=</td>
<td>1828</td>
<td>20</td>
<td>1270</td>
<td>22</td>
<td>432</td>
</tr>
</tbody>
</table>

Once again, this trend emerges even more clearly when we look at the relative frequency of epistemic SF within each context, shown in Table 3.26. Here we see once more a significant difference ($p \leq .0028$, $\chi^2 = 8.941664$) between temporal and epistemic SF, the
latter of which occur in 3% (17/616) of negative contexts and 6% (212/3746) of positive polarity contexts.

Table 3.26. Relative frequency of temporal and epistemic SF by polarity, all data sets

<table>
<thead>
<tr>
<th>Polarity</th>
<th>Temporal</th>
<th>Epistemic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Positive</td>
<td>3534</td>
<td>94</td>
<td>212</td>
</tr>
<tr>
<td>Negative</td>
<td>599</td>
<td>97</td>
<td>17</td>
</tr>
</tbody>
</table>

\[ p \leq .0028, \chi^2 = 8.941664 \]

At first glance, one may hypothesize that this difference may be due to a high relative frequency of epistemic SF in positive polarity interrogative contexts. An cross-tabulation of the relative frequency of temporal and epistemic SF reveals that the polarity effect is not due to interaction between polarity and sentence type: of the negative-polarity interrogatives, half (3/3) are SF with epistemic readings; of the positive-polarity interrogatives, 60% (18/30) are with epistemic SF. In declarative sentences, the SF constitutes 25% (9/36) of the negative polarity contexts and 21% (91/428) of the positive polarity contexts. At this time, the reasons for the difference in polarity tendencies between temporal and epistemic SF must be left unexplained.

### 3.6 Summary

This chapter has described the origins of the Spanish Synthetic Future and the ways in which both its form and functions have developed since the 13th century. In Section 3.1, I outlined the lexical origins of the Spanish SF since Late Latin. I further presented this construction as an example of a gram in the last stage of grammaticization according to Bybee, Pagliuca and Perkins' (1991) model for the development of modality-based futures, since it has developed an epistemic use. Section 3.2 showed that formal change has accompanied functional change, in that alternative variants that were contextually
restricted fell out of use, such that word order difference was no longer a viable syntactic option. We saw that before 17th century the SF showed marked distributional differences from the analytic form (e.g. cantar lo he), and subsequently, in the 17th century, from cantarélo. In Section 3.3, I addressed the decades-old problem of differentiating intention, prediction, and epistemic uses of the SF from a semantic perspective, with particular concern for empirical operationalization of semantic distinctions. In 3.3.1, I hypothesized that the semantic generalization of SF, which would entail a change in the contexts of use from intention to future prediction to epistemicity (Bybee & Pagliuca 1987), would be accompanied by i) a drop in first-person singular uses of SF; ii) an increase in the rate of co-occurrence of inanimate subjects with SF; and iii) a rise in the relative frequency of contexts involving prediction but not futurity. Regarding first-person singular distribution in SF, a significant drop in rates of co-occurrence was observed in temporal SF between the 19th- and 20th-century data sets, from 29% (124/432) to 20% (42/212), respectively (Table 3.6). At the same time, first-person singular rates are shown to be exceedingly low in epistemic SF, ranging from 2% (3/120) to 9% (4/43) of epistemic SF (Table 3.11). With respect to the second hypothesis, two periods of significant increase in the rate of co-occurrence of singular inanimate subjects with temporal SF were found: first, between the 13th- and 15th-century data sets (12% (130/1058) to 17% (135/770), respectively, and second, between the written and spoken 20th-century data sets (17% (37/212) to 25% (94), respectively), shown in Table 3.6. Table 3.11 shows that singular inanimate subjects are particularly likely to occur in contexts in which SF is understood epistemically, with rates of co-occurrence ranging from 21% (5/24) in the 19th-century data to 59% (13/22) in the 15th-century data. The
third hypothesis, that there would be a rise in the relative frequency of non-future-reference contexts of occurrence for SF, was found to be correct, as shown in Figure 3.4, with an increase from 2% relative frequency in the 19th-century Peninsular Spanish data to 24% in the 20th-century spoken data.

Sections 3.4 and 3.5 constitute the quantitative sections of this chapter. In Section 3.4, I describe the internal diachronic landscape of the Spanish SF both to refer to future events (Section 3.4.1) and used as a marker of epistemic modality (Section 3.4.2). The last subsection, 3.4.3, looks at the use of the SF as an imperative and to express "general truths" (Ultan 1978) within the larger context of future semantics. In Section 3.5, temporal and epistemic SF distribution patterns are compared, which reveals several contexts in which there are significant differences, suggesting that they should be considered, as Comrie suggested, separately in analyses, despite some semantic overlap (1985:47). It appears that the fundamental difference between epistemic SF and temporal SF is the fact that it almost never occurs in the first person (Table 3.11). A further difference is that the epistemic SF tends to be used to talk about states, and not actions, more often than the temporal SF (Table 3.12).

This chapter has revealed a future form whose patterns in future contexts have remained relatively the same over 700 years. This stability indicates that in many ways there has been little change in how the SF is used when it is used as a future, though changes in the distribution of subject suggest a move from intention meaning to prediction meaning. At this point, however, we have yet to delineate which, if any, of these patterns are particular to the SF construction. In order to begin to empirically define any special meaning the temporal SF itself may have, it will be necessary to compare the
patterns presented here with the patterns of other forms that perform closely related—perhaps even the same—functions of futurity. To this end, the next chapter will focus on the form that most closely overlaps with the temporal SF's function in modern Spanish: the Periphrastic Future (PF).
4 The Periphrastic Future

4.1 Origins

The Spanish Periphrastic Future (PF) is a periphrastic form made up of *ir* 'go' + *a* + infinitive (e.g. *voy a cantar* 'I'm going to sing'). Go-futures are common throughout the world's languages; along with Spanish, such futures have also been attested in Margi, Cocama, Maung, Atchin, Abipon, Krongo, Mano, English, Bari, Zuni and Nung (Bybee, Perkins & Pagliuca 1994:267). In Spanish, the PF is a more recent development than the SF; Fleischman (1982:82) hypothesizes that it probably did not gain conventional use as a future in spoken language until the 15th century (cf. Bersch in 1986:301, who places the emergence of the PF in the 16th century). Like the SF discussed in Chapter 3, the PF is hypothesized to have developed out of the conventionalization of inferences, this time those inferences common with purposive motion constructions. The inference most likely to co-occur in these constructions is that of intention. An often-cited example of the intention use with a go-future in English, in which a question about a locative goal is answered with a statement regarding intentions, is shown in (4.96).

(4.96)  
Duke Sir Valentine, whither away so fast?  
Val. Please it your grace, there is a messenger  
That stays to bear my letters to my friends,  
And I am going to deliver them.  
(c. 1595, Shakespeare, *Two Gentlemen of Verona* III.i.51, cited in Hopper & Traugott 1993:83)

---

66 In Old Spanish, this form often occurred without the intervening preposition *a*, as in *voy cantar* (Melis 2005).

67 Less common is the go-past, which survives, for example, in modern Catalan. See Pérez-Saldanya (2003) on this fascinating phenomenon in the history of Spanish.
Bybee, Perkins & Pagliuca (1994:280) suggest that, in the development of new futures, "intention is the crucial bridge to prediction and that the change from intention to prediction [i.e. future] occurs via the inferences that hearers make on the basis of the speaker's utterance." This change would be rather simple: "this use comes fairly directly from the literal meaning of 'the subject is on a path moving towards a goal'. The only necessary change is for the movement and path to be taken figuratively for the intention use to arise" (Bybee, Pagliuca & Perkins 1991:31).

Melis (2006:925-926, 933), however, disagrees with intention meaning as the semantic bridge in the case of the PF based on the fact that she claims to have found no first-person effect in her data: "lo que no muestran los datos es la interacción de la perífrasis de futuro con la intencionalidad. No vemos que los sujetos de primera persona—de los que cabe esperar una mayor capacidad volitiva-intencional—sean especialmente frecuentes" (Melis 2006:933).68 Instead, Melis argues for metaphorical reanalysis, with the speaker's perception of present relevance as the cognitive bridge. For example, citing Sweetser, she claims that, in the case of (4.97), "el punto de perspectiva está situado en el inicio del movimiento figurativo y establece el anclaje desde el cual el hablante visualiza el desarrollo gradual de la acción hacia su fin, como se mira a una persona alejarse del lugar donde se está en dirección a un punto de destino (Sweetser 1988:391-392)" (Melis 2006:925).

---

68 Nevertheless, Melis (2006:933n40) cites a rate of 50% (45/90) first-person subjects in early 19th-century PF usage. Though this is far from categorical first-person use of PF, this rate does seem elevated. Unfortunately, Melis offers no first-person SF proportion numbers, nor does she offer data for first-person PF in other time periods, for comparison.
¿Ya todos amamos? El mundo se va a perder (Celestina, VIII.213, cited in Melis 2006:925, 15c.)
'We all loved already? The world is going to be lost'

Melis' analysis is in line with Traugott's (1995:36, 50n4) proposal that the future meaning of the English be going to came about through "pragmatic reanalysis," i.e. through the inferential (re)alignment of the subject-experiencer of be going to and the speaker's perspective.

Whatever the intermediate step toward prediction may be, the earliest meaning of PF is undisputed. An early example, in (4.98), demonstrates the PF's proposed original meaning: "an agent on a path toward a goal" (Bybee & Pagliuca 1987).

estudia , mientras voy yo , a le decir tu pena (Celestina, Act 1, 15c.)
'study, while I go to tell (PF) him your sorrow'

The PF's proposed original meaning, then, is concrete and physical. The meaning "agent on a path toward a goal," is empirically verifiable (e.g. certain verb types, such as statives, would be incompatible with such a meaning), and will be used here as the springboard from which the development PF is analyzed.

4.2 Changes in form: Increased fixedness

It is difficult to determine at which point the PF became what may be considered a "grammaticized" form, i.e., when it was no longer simply a string of words, and instead a periphrastic construction. Hopper and Traugott (1993:122-123) offer the following rule of thumb:

One way of defining periphrasis is to characterize it as fulfilling the following criteria (cited in Vincent, forthcoming b [1991]; based on Dietrich 1973): the meaning of the periphrasis is not deducible from the constituent elements; the periphrastic construction shows syntactic unity at some level of analysis, where it did not do so before; the new periphrasis competes paradigmatically with other morphologically relevant categories.
Undoubtedly, unlike the SF in Chapter 3, the PF has not (yet) reached inflectional status, at least in written language. It is not uncommon, however, to encounter speculation that current trends point towards a syntheticization of the PF (e.g. Anderson 1979:22; Melis 2006:928n36); this, of course, would be expected as natural language change takes its course (Bybee, Pagliuca & Perkins 1991:34). There is evidence for such a change in certain American varieties, in which \textit{va a} is used for all persons/numbers, and functions like a prefix on the infinitive (Fleischman 1982:116; cf. Harris & Faarlund 2006:303).\footnote{These dialects are said to include colloquial educated speech of Panama City, varieties of rural Mexican Spanish, and a working-class variety of Salvadoran Spanish (Fleischman 1982:104).} At this point, however, given that the orthography used to represent the PF has not changed since the construction's inception, a quest for empirical evidence of increased fixedness in Peninsular Spanish—which is both suggested in the literature and a necessary step in grammaticization (e.g. Hopper & Traugott 1993:123)—becomes a bit trickier.

It is possible, nevertheless, to examine the degree of cohesion of the form. That is, how often does the form occur as a unit, and how often is it broken into parts due to intervening material, such as subjects or objects? In (4.99a), for example, the subject \textit{todo 'everything'} separates the conjugated \textit{ir} form from the \textit{a + INF}. Similarly, in (4.99b), both the subject \textit{usted 'you'} and the direct object \textit{ese escrito 'that writing'} intervene in the same position.
(4.99) **PF with intervening material**

(a) ¡Cé, hermano, que se va todo a perder! (Celestina, Act 4, 15c.)
'Hey, brother, everything is going to be lost (PF)!'  

(b) ¿Con que, en fin, va usted ese escrito a darme? (El afán, Act V, scene III, 19c.)
'So, finally, are you going to give me (PF) that writing?' (lit. 'going are you that writing to give me')

A decrease in such intervening material would indicate a move toward increased fixedness and increased automatization as a single processing unit (Bybee, Perkins & Pagliuca 1994; Torres Cacoullas 2001:445; Bybee 2003). Figure 4.5 shows the rates for intervening material with temporal uses of PF since the 15th century.  

**Figure 4.5. Rate of intervening material with temporal PF by data set**

As we can see, while a full third (33%, 4/12) of PF occurrences included intervening material in the 15th century, this rate dropped to 8% (5/59) in the 17th century, and to a mere 4% (3/75) by the 19th century. This possibility no longer occurred in these data in 20th-century writing, and it made up only 1% (8/768) of 20th-century speech (cf. Villa-Crésap 1997:63, who also suggests that this phenomenon has a very low frequency). This
suggests that the temporal PF, at least in terms of fixedness, acquired full grammatical status in the 20th century, a full 300 years after the SF passed the same milestone (see Section 2.2). This moment, however, is not properly a "moment," for this change was not sudden or instantaneous; thus, the "becoming" of PF as a construction, as with the SF, cannot be pinpointed diachronically, but rather can only be referred to in relative terms.\(^{70}\)

4.3 The functions of PF

4.3.1 Temporal PF

The temporal PF, in contrast to the SF, is rarely described on its own (but see Melis 2006), despite the fact that it has become the most frequent future form in many varieties of American Spanish (Cartagena 1995-1996): in fact, several have claimed that it is actually the default (e.g. Sedano 1994:238; Silva-Corvalán 1994:22; Balestra 2006). Where the PF has received its share of attention is in the work of linguists who have wished to identify where, if ever, its use overlaps with the SF, and how speakers make the choice (e.g. Kahane & Hutter 1953; Reid 1955; Montes 1962-1963; Gili Gaya 1964; Imbs 1968; Sáez Godoy 1968; Paufler 1970; Lope Blanch 1983 [1972]; Tlaskal 1978; Fleischman 1982; Berschin 1986:303; Gómez Torrego 1988:66; Villa-Crésap 1997; Fernández de Castro 1999:206; Llorente Vigil 1999:25-27). As described in Section 1.1.2.1, the PF has been associated with a variety of meanings, most often linked to its lexical allative origins. Vet (1994), for instance, argues that the PF signals that the speaker perceives the preparatory state for the action to be complete, as in (4.100), as

\(^{70}\) Another criterion for measuring fixedness would be clitic climbing (Myhill 1988). Unfortunately, I did not code for clitic climbing in this study, and thus such an examination has been left for subsequent research.
indicated by *según las señas* 'according to the signs'. For a more in-depth look at scholars' hypotheses on PF semantics, please refer to Section 1.1.2.1.

(4.100) ¿Tú *vas a salir*, según las señas? (El afán, Act II, Scene III, line 52, 19c.)
'You're going to go out (PF), according to the signs?'

One of the reasons behind the lack of studies focusing solely on the PF is that the PF is a much younger form and, in turn, much less complex both in form and function. It has not, in fact, changed its form in any notable way since its beginnings, except in increased fixedness, shown in Section 4.2, or perhaps in as of yet unstudied phonological reduction. Cartagena (1995-1996:95) also points out that the PF is rarely mentioned as a future tense in traditional grammars, a trend he sees as evidence of grammarians' bias toward recognizing as tenses only those forms that developed from constructions already present in Latin (cf. Poplack & Malvar forthcoming for reports of a similar lack of acknowledgement of PF in Portuguese grammars).

A further reason may be that the semantics of PF are much less contentious than those of the SF; previous studies have focused on the semantic shift from a purposive motion construction to a future construction on the one hand (e.g. Bybee & Pagliuca 1987; Bybee, Perkins & Pagliuca 1994; Melis 2006), and/or its competition with SF and other constructions as a marker of futurity, on the other (e.g. Poplack & Turpin 1999; Aaron 2006; Poplack & Malvar forthcoming). As Bybee, Pagliuca and Perkins (1991:31) note, "movement-derived futures do not provide the richness of semantic nuance that modality-derived [such as the SF] futures do." I will return to the temporal PF in Section 4.3.1, with evidence about its grammaticization. Before this is possible, however, it is necessary and helpful to look more closely at the rarely mentioned—perhaps rightly so—"other" uses of the Spanish PF.
4.3.2 Other proposed PF functions

Unlike the SF, which has been assigned a "plethora of meanings" (Poplack & Malvar forthcoming; see Chapter 1), the possible polysemous nature of the PF has been, on the whole, ignored. Instead, most scholars have focused on differentiating it semantically from the SF (see Chapter 4). Nonetheless, a very few scholars have noted that the PF often occurs in construction with an epistemic modal interrogative, as in (4.101) (e.g. Tlaskal 1978; Bauhr 1989; Sedano 1994; Cartagena 1995-1996; Villa-Crésap 1997; Melis 2006). This construction is rhetorical and implies incredulity (Melis 2006:921). The few scholars who have mentioned it have gone so far as to label the use of PF in this construction a different (non-future) use of the PF. For example, Cartagena (1995-1996:96) refers to this as the "futuro replicativo" ('replicative future') and Sedano (1994:228) as "rhetorical" periphrasis (see also Bauhr 1989:133-135, as well as Tlaskal 1978:209 on Portuguese).

(4.101) Dice: "Tú me vas a decir dónde está Juana," dice. "¿Cómo voy a pensar yo que no... no vas a saber, si tú andas con este hombre?", le dijo. (MexPop 194)

'He says: 'You are going to tell me where Juana is,' he says. 'How am I going to think/am I supposed to think (PF) that you aren't going to know/wouldn't know, if you're going out with this guy?' he told her.'

In these examples, however, we find that the PF has little to do with a (posited) meaning of rhetorical incredulity; it simply contributes futurity. A strong piece of evidence that this construction's meaning is not dependent upon the PF is that the SF also appeared in this construction, as shown in (4.102)-(4.106).\(^{71}\)

\(^{71}\) Note that (4.103) is also what has been termed a "general truth" use, i.e., it occurs with a nonspecific subject.
(4.102) Et dixeron el lobo et el cuervo: ¿Cómo conbrá a ti, ca tú sabes que hueles muy mal et has el vientre líxoso? (Calila, El camello que se ofreció al león, p. 162, 13c.)

'And the wolf and the crow said: And how could he eat (SF) you, because you know that you smell very bad and your intestines are disgusting?'

(4.103) La mayor desconociencia es quien non conosle a sís; pues ¿cómo conozcá a otrí? (Lucanor, Segunda parte, p. 283, 14c.)

'The greatest ignorance is he who knows not himself; because how will he know (SF) another?'

(4.104) Et yo ¿, dó fuiré de ti o dó me apartaré? ; et deviate esquivar. ¿, Et cómo avrán tu esperança en tu lealtad, aviendo tú fecho esto a tu rey que te onrara quanto yo vi? (Calila, La garza, la culebra y el cangrejo, p. 175, 13c.)

'And how will they have (SF) your hope in your loyalty, you having done this to your king who honored you as far as I could see?'

(4.105) Dixo ella: - ¿ Et cómo non seré triste, que la mi vida non era de ál sinon de las ranas, et agora vinome grand ocasión, de guisa que non puedo comer nin tomar sinon las que me dan en limosna? (Calila, La culebra y las ranas, p. 249, 13c.)

'She said: And how could I not be (SF) sad, for my life was not anything but that of frogs, and now a great occasion came to me, such that I cannot eat or drink but what they give to me in charity?'

(4.106) si al tal Sancho Pança le das insula que gouierne, porque el que no sabe gouernarse a si, ¿cómo sabré gouernar a otros? (Quixote, Cap. XXXIII, fol. 129r, 17c.)

'If to this Sancho Panza you give an island to govern, because he doesn't know how to govern himself, how will he know (SF) how to govern others?'

As we can see, then, the special meaning of this construction does not depend on the meaning of the PF, which is simply doing its semantic work of prediction, but in a specialized context.72 For example, (4.102) would mean 'how do you predict that he would eat you?', and (4.103) would mean 'how do you predict that he would know

---

72 Given the fact that the use of SF in this context is not found in these data after the 19th century, that is, after the birth of the PF into futurity, it may be quite accurate to insist on an association of the PF with this particular construction, i.e. speakers tend to use the PF in this construction more often than the SF; nonetheless, it is the construction, not the PF itself, in conjunction with the prediction meaning of PF, that expresses incredulity.
another?" The incredulity meaning may be contributed by various elements in this construction, including, perhaps, question words (which often contribute epistemicity, Narrog 2005:679-680) and intonation, though these contributions have not yet been examined in the literature on the Spanish PF.

Another use of the PF sometimes mentioned is the exhortative (PF) (Montes 1962-1963; Sedano 1994:229; Villa 1997), as in (4.107), which, of course, occurs in the first-person plural.

(4.107) Oye **vamos a subir** arriba que estamos aquí de pie... (COREC, CCCON022B, 20s)
'Hey we're gonna go upstairs/let's go upstairs (PF) since we're here standing'

The exhortative in PF, however, like the imperative in SF, cannot always be distinguished from other temporal uses, since both involve reference to a future event; an imperative reading relies upon the occurrence of a suitable subject (i.e. second-person for imperative, first-person plural for exhortative) and the perception of speaker assertion (see Section 3.4.3 for exemplified justification of this position). As further evidence of the inextricability of futurity and these imperative uses, I would like to emphasize two points. First, the PF not only functions as an exhortative/imperative in the first-person plural, but it can also be interpreted as an imperative with other persons. Let us look at (4.108). Here, the speaker seems to be making a demand of his interlocutor based on his (assumed) previous knowledge about her ability to do what he is demanding.

(4.108) Dice: "Tú me **vas a decir** dónde está Juana," dice. "¿Cómo voy a pensar yo que no...no vas a saber, si tú andas con este hombre?", le dijo. (MexPop 1994)
'He says: "You are going to tell (PF) me where Juana is," he says. "How am I supposed to think that you wouldn't know, if you're going out with this guy?" he told her.'
Second, it is important to note that both SF and PF are used in these same (apparently imperative) contexts, suggesting a strong semantic tie between prediction and exigency, making this a characteristic of the socio-pragmatic realm of futurity, not that of any one particular form (cf. Bybee, Perkins & Pagliuca 1994:210-211).

4.3.3 Non-temporal PF

The only use of PF that is undoubtedly not future is its lexical allative meaning in certain contexts (Berschin 1986:301). This use involves a purposeful going, most often habitual, as in (4.109), which co-occurs with the temporal adverbial los domingos 'on Sundays', indicating habituality.

(4.109) que los domingos va la familia a hacer la compra. (COREC, CCCON034B, 20s)

'that on Sundays the family goes to shop (PF)'

It can also appear—especially in data from plays—in stage directions, as in (4.110), or historical present, as in (4.111), both of which describe an event as it is occurring.

(4.110) MICHU abandona el coche con mil precauciones y va a colocarse detrás de su padre, es evidente que ve en él un buen escudo. (Billy, p. 74, 20w)

'MICHU abandons the car very cautiously and he goes to stand (PF) behind his father, it's obvious that he sees him as a good shield'

(4.111) y yo estaba con la túnica y estaba bailando, como parece un camisón, yo ahí dando vueltas, todos cantando y yo bailando y así. Pues de repente dijo: 'Estoy ya muy aburrido, me voy a echar a volar un rato'. Y me asomo, me voy a asomar a una ventana, pero resulta que es que era la ventana del cuarto de baño del otro piso. (COREC, CCCON032A, 20s)

'and I was with the dress and I was dancing, since it's like a robe, me there spinning around, everyone singing and me dancing and like that. Well suddenly he said: 'I'm really bored already, I'm gonna go wander around for a while'. And I walk up, I go to walk up (PF) to a window, but it turns out that the thing is that it was the window for the bathroom on the other floor.'

The existence of this use is a clear example of retention (Bybee & Pagliuca 1987:212), in which older uses of a construction co-exist alongside more innovative uses. In this case, it
would appear that divergence is manifest in polysemy, such that some uses mean prediction, while others are purposive motion constructions with no future meaning. While the latter do exist, and polysemy is a characteristic of PF semantics, these non-future uses are not the premier site where retention of early PF meanings is found. Let us look at the rate of PF constructions with no possibility for temporal readings, shown by century in Figure 4.5.

Table 4.27. Occurrence of non-future motion PF within all PF by data set

<table>
<thead>
<tr>
<th>Data set</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>OldSp</td>
<td>3/16</td>
<td>19</td>
</tr>
<tr>
<td>17</td>
<td>0/59</td>
<td>--</td>
</tr>
<tr>
<td>19</td>
<td>2/77</td>
<td>2</td>
</tr>
<tr>
<td>20w</td>
<td>1/80</td>
<td>1</td>
</tr>
<tr>
<td>20s</td>
<td>7/775</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

As we can see, the presence of PF constructions that do not allow a temporal or prediction reading has been negligible since the inception of this construction, most often between 0% and 2%. Old Spanish data show an uncharacteristic 19% (3/16), but when PF token frequency is appreciable, after the 17th century, there are very few pure motion (i.e. non-future) tokens. These low rates suggest that evidence for the retention of earlier PF meanings may best be found within the realm of future expression, and not in PF divergence.

4.4 The grammaticization of the PF

As mentioned earlier, Fleischman (1982:82) estimates that the PF was conventionalized as a future in spoken language by the 15th century. Recently analyzed textual evidence

73 These findings contrast with those of Poplack and Turpin (1999:160n14) for Canadian French, in which they found that the PF "is used nearly as often for non-future reference, largely in habitual contexts."
has shown that the PF began to be used as a future sometime between the 17th century (Aaron 2006; and present study, Section 4.2) and the 19th century (Melis 2006:930), depending on corpus size and the analysts' interpretations of what constitutes "conventionalization." My own estimates, if we assume a lag time of about 150 years between spoken and written language, are in line with Fleishman's, while Melis' (2006) estimate would push forward Fleishman's estimate about 200 years. These estimates must be taken, obviously, as extremely rough at best. Biber (1995:Ch. 8) shows that the distance between speech and writing is not steady over time, and, furthermore, that there are significant differences between different registers, such that, according to various measurements, some written genres approximate speech more than others, and certain registers of speech approximate writing more than others. Vast diatopic differences in the progression of this change have also been noted by several scholars (e.g. Cartagena 1995-1996; Westmoreland 1997; Makihara forthcoming). Though the hypothesized chronology for PF use as a future in my data are in synch with Fleishman's and Melis' estimates, since the PF achieves a relative frequency of 12% of SF-PF futures in the 19th century (compared to a mere 4% in the 17th; see Chapter 5), it is not, in fact, that the PF began to be used as a future in the 19th century, having developed from an older, more lexical meaning; it is more correct to say that the PF began to be used as a future in the 19th century, i.e. it

---

74 Melis (2006:931) found that the PF (in relation to the SF) had a relative frequency of 2% in the 18th century, 5% in the early 19th century, 16% in the late 19th century, and 30% in the early 20th century in a corpus that included both Peninsular and Mexican texts, and which included all occurrences of both PF and SF apparently irrespective of meaning (see Melis 2006:929). Whether a rate of 5% in the 19th century in Melis' study, or 4% in the 17th century and 12% in the 19th century in this study, should be considered an indicator of "conventionalized" status is debatable.
came of age as a construction and *began to matter* in quantitative terms, reaching over 10% relative frequency (alongside the SF) in Peninsular Spanish.

The key to the success of this construction as a future marker may lie in this fact: its meaning almost always allowed for a temporal, future interpretation (unlike SF, which has been characterized as having passed through obligation/predestination, though such uses were not found in these data). I cannot claim here that this potential was always exploited by hearers or that this was the most salient of the semantic possibilities, any more than I can claim that it was not. Melis (2006:934, emphasis added) notes:

> Es preciso aclarar que entre la percepción objetiva del inicio de un proceso (*ir* = movimiento temporal) y la representación mental del futuro que emerge del presente (*ir* = movimiento abstracto) se dibuja una frontera sumamente fluida. Esta es la razón por la cual el deslizamiento del valor inminente al valor prospectivo tiene manifestaciones *desde el principio de la historia*, según documentan todos los casos de *ir a* + infinitivo procedentes del siglo XIX que aparecieron marcados con el rasgo [-inminente].... Estos casos ponen de manifiesto que la *perifrasis es aprovechada de inmediato para los fines del mensaje subjetivo*, con implicaciones, claro está, de cierto traslape entre las dos primeras fases del cambio.

If we look back at (4.96), repeated here for convenience, we find multiple possible points of semantic salience: motion, intention and future. It is impossible to tell for sure which of these would be most salient for a hearer, especially a hearer of the late 16th century.

(4.96) *Duke*  Sir Valentine, whither away so fast?

*Val.*  Please it your grace, there is a messenger
That stays to bear my letters to my friends,
And I am going to deliver them.

(c. 1595, Shakespeare, *Two Gentlemen of Verona* III.i.51, cited in Hopper & Traugott 1993:83)

Of primary concern in the development of PF, then, is not *what* it could mean, but rather through *where* it could mean it, i.e., the contexts in which it occurred. If grammaticization has indeed occurred, it may be best observed in changes in the
contextual distribution of a form or construction over time. It is in these patterns that we may verify contextual generalization or semantic bleaching (Torres Cacoullos 2000:71).\footnote{This was the case, for example, in the similar case of Spanish progressives \textit{estar} + \textit{n]do}, \textit{ir} + \textit{n]do} and \textit{andar} + \textit{n]do}, which showed a diachronic decline in co-occurring locative adverbials as their motion meaning was bleached (Torres Cacoullos 2000:Ch. 3).} We would posit, then, that the PF was not becoming a future, but rather a better (i.e. more contextually and semantically generalized) future. In the case of the PF, we will see evidence for contextual generalization in three sites in particular. First, we will see that, while early PF was limited to particular verb classes, those harmonic with its lexical origins, the spoken 20\textsuperscript{th}-century PF is permitted with all verb classes, including those that are incompatible with motion meaning, such as stative and psychological verbs. Furthermore, we will see that the type-token frequency of PF has increased, and that this construction's rate of co-occurrence with locative adverbials has decreased, demonstrating a loss of locative meaning.

4.4.1 Verb class

Verb class patterns with PF show that the early PF, the newborn future, was also a very limited future. In Table 4.28, we see that the PF could not occur with all verb classes until 20\textsuperscript{th}-century speech. In Old Spanish, the temporal PF occurred in dynamic verbs, both motion at 15\% (2/13) and non-motion at 69\% (9/13), and verbs of perception at 15\% (2/13). In the 17\textsuperscript{th} century, the only verb that did not occur with the PF was \textit{ir} 'go', a restriction that was maintained in all written data. Still, dynamic and motion verbs make up over 80\% of the PF occurrences, at 64\% (38/59) and 17\% (10/59), respectively. This dominance remained constant in all written data sets: in the 19\textsuperscript{th} century they account for
83% (61/75) of temporal PF occurrences; and in the 20th-century written data they make up 81% (63/79). Only in the 20th-century speech data do we see the dominance of dynamic verbs weakening; here, they make up 66% (504/768) of PF occurrences, and stative verbs, after 600 years of representing a mere 8% or less of PF occurrences, now account for 20% (155/768). Also, for the first time in these Peninsular data, ir 'go' occurs in construction with PF, making up 6% (47/768) of temporal PF uses in the speech in this corpus.

Table 4.28. Distribution of verb class in temporal PF by data set

<table>
<thead>
<tr>
<th>Verb class</th>
<th>OldSp % (N)</th>
<th>17 % (N)</th>
<th>19 % (N)</th>
<th>20w % (N)</th>
<th>20s % (N)</th>
<th>Total % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic</td>
<td>69 (9)</td>
<td>64 (38)</td>
<td>69 (52)</td>
<td>72 (56)</td>
<td>58 (444)</td>
<td>60 (599)</td>
</tr>
<tr>
<td>Stative</td>
<td>--</td>
<td>5 (3)</td>
<td>8 (6)</td>
<td>8 (6)</td>
<td>20 (155)</td>
<td>17 (170)</td>
</tr>
<tr>
<td>Motion</td>
<td>15 (2)</td>
<td>17 (10)</td>
<td>14 (9)</td>
<td>9 (7)</td>
<td>8 (60)</td>
<td>9 (88)</td>
</tr>
<tr>
<td>Ir</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>6 (47)</td>
<td>5 (47)</td>
</tr>
<tr>
<td>Perception</td>
<td>15 (2)</td>
<td>10 (6)</td>
<td>4 (3)</td>
<td>2 (2)</td>
<td>4 (30)</td>
<td>4 (43)</td>
</tr>
<tr>
<td>Psychological</td>
<td>--</td>
<td>3 (2)</td>
<td>7 (5)</td>
<td>9 (8)</td>
<td>4 (32)</td>
<td>5 (47)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total PF N</td>
<td>13</td>
<td>59</td>
<td>75</td>
<td>79</td>
<td>768</td>
<td>994</td>
</tr>
</tbody>
</table>

As Table 4.28 shows, dynamic verbs as in (4.112) were the most robust context in which speakers first made use of this construction. Motion verbs as in (4.113) also showed a strong presence in early PF. A tendency to occur with motion verbs (excluding ir 'go') in a motion-based construction like PF makes sense, since the already-present ir 'go' is semantically harmonic—yet not completely redundant—with other motion verbs (see J. Lyons 1977 on semantic harmony; cf. Bybee & Dahl 1989; Krongauz 1998, Ch.6). These motions verbs, far from clashing with the semantics of ir,

---

76 This finding is different from Poplack and Tagliamonte's (2001:227) results for be going to in mainstream Ottawa English, which showed no statistical significance of verb class, with verbs of motion having a Probability of [.51] and other verbs having a Probability of [.50].
support it and allow for other inferences common with the PF construction to become more salient. Motion verbs most frequently occurring in the data are venir 'come', salir 'go out', volver 'return', llegar 'arrive', and entrar 'entrar'. For instance, in (4.113), a 20th-century written example that clearly displays motion meaning, as evidenced by the co-occurrence of two locatives: de dónde 'from where' and del cañón 'from the cannon', the verb salir 'go out', contributes motion meaning to the expression. This contribution makes the motion interpretation less dependent on the verb ir 'go', potentially allowing secondary meanings, such as intention and posteriority, to become more salient.

(4.112) el tal hombre jura que **va a morir** en la horca, y si muere en ella juró verdad, y por la ley puesta merece ser libre, y que passe la puente; y si no le ahorcian, juró mentira, y por la misma ley merece que le ahorquen. (Quixote, Cap. LI, fol. 194v, 17c.)

'this man swears that he is going to die (PF) at the noose, and if he dies there, he swore the truth, and by law he deserves to be free and cross the bridge; and if they don't hang him, he swore a lie, and by the same law he deserves to be hanged.'

(4.113) ya me contarás tú **de dónde van a salir** húsusmos del cañón de mi Santa Bárbara (Pobres diablos, p. 92, 20w)

'you tell me where "obusismos" are going to come out (PF) from the cannon of my Saint Barabara'77

On the other hand, the use of the PF in stative (4.114) and psychological (4.115) contexts was limited, though progressively less so; use of the SF in such contexts, as in (4.116), was still more common (see discussion of Table 5.45). Most impressively, the verb ir 'go', shown in (4.117), simply did not occur with PF in any of the written data, appearing only in 20th-century speech. The reason behind speaker avoidance of such phrases as voy a ir 'I'm going to go' is most likely evidence of retention of allative meaning for voy, thus making such a statement fully redundant (unlike other motion verbs, which contribute

77 The meaning of "obusismos" is not clear. This word is not listed in the Diccionario de la Real Academia, and an internet search for this word returned no results.
more semantic content than simple going) and unacceptable. The occurrence of *ir* 'go'
with PF would indicate that the allative meaning had been sufficiently bleached for
speakers not to perceive this as unnecessary redundancy.

(4.114) Padre mío, ¿qué va a ser, en faltándole tú, qué *ya a ser* de tu hija!... (Conjuración,
Act II, Scene III, 19c.)
'My father, what will be[come], with you gone, what will be[come] (PF) of your
daughter!'

(4.115) Hoy mismo, hoy mismo *ya a saber* todo el mundo que soy esposa de Rugiero!
(Conjuración, Act V, Scene VI, 19c.)
'today, this very day everyone is going to know (PF) that I am Rugerio's wife!'

(4.116) Infeliz Laura, ¿qué *será* de ti?... Un presentimiento fatal me estrecha el corazón y
ni me deja respirar siquiera... (Conjuración, Act II, Scene II, 19c.)
'Miserable Laura, what will be[come] (SF) of you?... A horrible feeling squeezes
my heart and hardly lets me breathe'

(4.117) Que - bueno no sé - el - el jueves si *vas a ir* o yo *voy a ir*. (COREC,
CPCON006A, 20s)
'(That) – well I don't know – on – on Thursday if you are going to go (PF) or I am
going to go (PF).'</n
The generalization of PF evidenced by its emergence in contexts incompatible with
motion (Table 4.28), such as stativity (e.g. Traugott 2003:365), shows that the PF
acquired increasing functionality as speakers began to use it to perform the
communicative task of general prediction, no longer necessarily linked to motion.

### 4.4.2 Lexical diffusion

The tendency of the PF to occur with certain verb classes raises the question of lexical
diffusion (see, e.g., Kiparsky 1982; Tottie 1991; Cole & Hualde 1998): is the distribution
of the PF with certain verb classes (Table 4.28) in fact attributable to certain lexical
types? Lexical effects have been found in the use of Romance futures before; these
effects, however, have been linked to conservative forms that are being replaced. Poplack
and Turpin (1999:156), for instance, found that the SF in Canadian French tended to occur with certain verbs: vouloir 'want', savoir 'know', revenir 'return', être 'be', and avoir 'have'; they associate these lexical effects with the non-productivity of this construction, and further assert that these lexical effects override any verb class effects that may have once been salient. Aaron (2003) also associated lexical effects of decir 'say' and ver 'see' in Mexican Spanish SF with a lack of productivity. It would be quite possible, however, to find lexical effects at the other end of the path as well, i.e., the beginning. This was the case, for example, in do-support in English, for which Ellegård (1953:201) states, "It is in fact not unlikely that each verb has its own history," as well as in English negation (Tottie 1991).

In the case of the PF, then, my aim here is to discover if the tendency to occur with particular verb classes seen above in Section 4.4.1 is, in fact, a tendency to occur only with particular verbs. If this is the case, we would expect that these highly frequent verbs would persist as most frequent among PF uses for as long as these lexical effects were salient in PF use. Table 4.29 shows the five most frequent lexical types in PF by dataset, in descending order of frequency, with the number of tokens shown in parentheses. For Old Spanish I only list the three that occurred more than once.78

---

78 This table (Table 4.27) includes both temporal and motion PF (see Table 4.1). If motion PF were removed, the only change in the lists would be that one token of Old Spanish decir 'say' would not be counted, and decir would be removed from the Old Spanish list.
Table 4.29. *Most frequently occurring lexical types in all PF by data set*\(^\text{79}\)

<table>
<thead>
<tr>
<th>OldSp</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td>decir (2)</td>
<td>buscar (14)</td>
<td>decir (7)</td>
<td>contar (6)</td>
<td>hacer (66)</td>
</tr>
<tr>
<td>perder (2)</td>
<td>decir (6)</td>
<td>dar (6)</td>
<td>hacer (6)</td>
<td>ir (47)</td>
</tr>
<tr>
<td>ver (2)</td>
<td>ver (6)</td>
<td>hacer (6)</td>
<td>pasar (3)</td>
<td>ser (43)</td>
</tr>
<tr>
<td>parar (3)</td>
<td>ser (4)</td>
<td>convertir (3)</td>
<td>decir (28)</td>
<td></td>
</tr>
<tr>
<td>acompañar (2)</td>
<td>ver (3)</td>
<td>decir (2)</td>
<td>ver (28)</td>
<td></td>
</tr>
</tbody>
</table>

Total % (N) =

38 (6/16) 52 (31/59) 34 (26/77) 25 (20/80) 27 (212/775)

Table 4.29 offers two pieces of information of interest in the generalization of the PF.

First, we see that no particular lexical items have dominated PF occurrence; *decir* 'say' is the only type that appears as one of the five most frequent in all data sets, though not necessarily at the top. Besides the early tendency to co-occur with dynamic and motion verbs, no lexical item appears particularly salient.\(^\text{80}\)

The second piece of information offered by Table 4.29 is the percentage of the data set's tokens represented by the verbs listed for that century. For example, the three

\(^\text{79}\) Table 4.27 includes the following verbs: *acompañar* 'accompany', *buscar* 'look for/get', *contar* 'tell', *convertir* 'change into', *dar* 'give', *decir* 'say', *hacer* 'make/do', *ir* 'go', *parar* 'stop', *pasar* 'happen/pass', *perder* 'lose', *ser* 'be', and *ver* 'see'.

\(^\text{80}\) At first glance, one is tempted to say that the PF has been particularly associated with the verbs *decir* 'say', which is of the five most frequent types in all data sets, and *ver* 'see', which occurs in the most frequent lists in all data sets except the 20\(^\text{th}\)-century written data. This is, indeed, the case. It is also the case, however, that the SF may also be associated with these same verbs, given the fact that these verbs form the fixed phrases *Qué te diré* 'what will I tell you' and *Ya verás* 'you'll see' in SF in Mexican Spanish (Aaron 2003); the latter occurs in these Peninsular data as well (see Section 2.2). *Decir* 'say' is highly frequent in other (i.e. non-future) contexts as well, as demonstrated in Travis' (2005) work on epistemic uses of this verb in conversational Spanish.
verbs listed as most frequent in Old Spanish PF account for 38% (6/16) of the PF data. This rate is 52% (31/59) for the 17th century, 34% (26/77) for the 19th, 25% (20/80) for 20th-century written data, and 27% (212/775) for 20th-century spoken Peninsular Spanish. Thus, we see a decline in the dominance of the most prominent verbs in PF. Another way to explore the possible diversification of PF in terms of lexical co-occurrence is to calculate the type-token ratio. These calculations are seen in Table 4.30.

Table 4.30. *Type frequency and type-token ratio in all PF by data set*

<table>
<thead>
<tr>
<th></th>
<th>OldSp</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type frequency</td>
<td>13</td>
<td>30</td>
<td>47</td>
<td>59</td>
<td>212</td>
</tr>
<tr>
<td>Type-token ratio</td>
<td>.76</td>
<td>.51</td>
<td>.61</td>
<td>.74</td>
<td>.27</td>
</tr>
<tr>
<td>Total N=</td>
<td>17</td>
<td>59</td>
<td>77</td>
<td>83</td>
<td>830</td>
</tr>
</tbody>
</table>

Here we see that along with PF token frequency, the type frequency in PF has risen, from 13 in Old Spanish, 30 in the 17th century, 47 in the 19th century, 59 in 20th-century literature and 212 in 20th-century speech. The type-token ratio in the PF shows the productivity of the PF; the higher the type-token ratio, the more productive the form. Surprisingly, this ratio is highest in the Old Spanish data, at 81%. The 17th century shows a ratio of 51%, which rises to 61% in the 19th century and 71% in 20th-century literature. In written genres, then, we see an increase in overall PF productivity between the 17th and 20th centuries. The spoken data show a very different picture: the type-token ratio is only 27. It is likely that this lower diversity rate reflects not of a new PF tendency to co-occur with certain lexical items, but rather a general difference between speech and writing: natural speech, which is produced on-line and is thus unplanned, uses a different and

---

81 If this table had included two more verbs, both occurring once in the data, this percentage would have been 50% (8/16).
narrower vocabulary than writing, which is pondered and often edited (cf. Biber 1995:283 on planning and editing of texts, and Biber 1995:292-293 on the association of high type-token ratio with written registers). In general, then, a rise in the type-token ratio points to a general diachronic lexical diffusion of PF.

4.4.3 Subject

In Chapter 1, I discussed the association of grammatical person and subject animacy with certain meanings that have been associated with the PF in the literature, such as volition and intention. It was noted that person and animacy can give some clues into these meanings; intention is best expressed in first person singular (e.g. Sedano 1994; Melis 2006), and volition by animate subjects (e.g. Bybee & Pagliuca 1987). According to the diachronic path for grammaticization of go-futures proposed by Bybee, Pagliuca and Perkins (1991; cf. Bybee, Perkins & Pagliuca 1994:280), the PF should have had an increase in intention meaning sometime early in its emergence as a competing future form. This intention meaning would subsequently fade out as the PF came to express general prediction, which they consider to be the essential meaning of "future." If this is the case, then we would expect to see an increase in first-person uses at the time that we have posited the PF to be conventionalized as a future form, i.e., the 19th century (Section 4.4, cf. Fleischman 1982; Melis 2006). We would further posit that, as the PF followed its path from intention to prediction, there would be a decrease in first-person uses.

In Table 4.31, we see the distribution of subjects in temporal PF occurrences in each data set. Focusing on first-person singular uses, we see that the lowest rate occurred in Old Spanish, at 23% (3/13), followed by the 17th century, at 27% (16/59), when the PF accounted for only 4% of PF/SF uses in future contexts (see Section 5.2). In the 19th
century, when scholars (e.g. Fleischman 1982; Melis 2006), including myself, have found that the PF reached a considerable relative frequency (12% in these data, see Section 4.4), the proportion of first-person uses, like in (4.118), which shows a request for and expression of intention, increases to 40% (30/75).

(4.118) ...ni a qué has venido, ni...
SIMÓN.-- Bien, a eso voy... Sí señor, **voy a decir** a usted... Conque... Pues el amo me dijo... (El sí, Act II, Scene X, 18-19c.)

'...nor what you have come to do, nor... SIMÓN: Well, that's what I'm getting to. Yes sir, I'm going to tell (PF) you... That... Well the boss told me...'

The difference between this 19th-century rate and the first-person singular rates for previous data sets approaches but does not reach significance at the .05 level ($p \leq .0801$; $x^2 = 3.0625$), most likely due to the low token frequency of these early stages. After the 19th century, we see a steady decline in first-person uses, to 35% (28/79) in 20th-century written texts and 30% (229/768) in 20th-century speech. Though these do not constitute statistically significant changes, they do suggest a decline in intention use of PF after the 19th century. These results are further corroborated by Melis’ (2006:933n40) findings for the early 19th century, which in her data had a first-person rate of 50% (45/90) (presumably including both singular and plural, though this is not specified), a figure she cites as evidence that these first-person subjects were not particularly frequent (Melis 2006:933). If we compare these numbers to other data, however, we see that a rate of 50%, in Melis’ data, or 40% here, for the 19th century, does indeed show that the first person singular was particularly frequent, since the SF shows an average rate of 30% (1241/4121) first-person singular subjects and the average for the PF for all data sets is 31% (306/994). This suggests that intention meaning, as, for example, Bybee, Pagliuca

146
and Perkins (1991) suggest, did in fact serve as one of the key meanings between motion and prediction, though the presence of inanimate subjects with the PF since Old Spanish, as in (4.119), with the subject todo 'everything', suggests that the prediction meaning was always possible.

(4.119) ¡ Ce, hermano, que se va todo a perder! (Celestina, Act 4, 15c.) 'Oh, brother, (that) everything is going to be lost (PF)!

Table 4.31. Distribution of subject in temporal PF by data set

<table>
<thead>
<tr>
<th>Subject</th>
<th>OldSp</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
</tr>
<tr>
<td>1s</td>
<td>23 (3)</td>
<td>27 (16)</td>
<td>40 (30)</td>
<td>35 (28)</td>
<td>30 (229)</td>
<td>31 (306)</td>
</tr>
<tr>
<td>2s</td>
<td>15 (2)</td>
<td>17 (10)</td>
<td>12 (9)</td>
<td>20 (16)</td>
<td>14 (109)</td>
<td>15 (146)</td>
</tr>
<tr>
<td>3s, human</td>
<td>15 (2)</td>
<td>19 (11)</td>
<td>16 (12)</td>
<td>11 (9)</td>
<td>14 (110)</td>
<td>14 (144)</td>
</tr>
<tr>
<td>3s, inanimate</td>
<td>15 (2)</td>
<td>--</td>
<td>17 (13)</td>
<td>10 (8)</td>
<td>12 (93)</td>
<td>12 (116)</td>
</tr>
<tr>
<td>Animals, s&amp;p</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>&lt;1 (3)</td>
<td>&lt;1 (3)</td>
<td>--</td>
</tr>
<tr>
<td>1p</td>
<td>23 (3)</td>
<td>29 (17)</td>
<td>5 (4)</td>
<td>10 (8)</td>
<td>18 (138)</td>
<td>17 (170)</td>
</tr>
<tr>
<td>2p</td>
<td>--</td>
<td>2 (1)</td>
<td>4 (3)</td>
<td>4 (3)</td>
<td>2 (13)</td>
<td>2 (20)</td>
</tr>
<tr>
<td>3p, human</td>
<td>8 (1)</td>
<td>7 (4)</td>
<td>3 (2)</td>
<td>8 (6)</td>
<td>5 (39)</td>
<td>5 (52)</td>
</tr>
<tr>
<td>3p, inanimate</td>
<td>--</td>
<td>--</td>
<td>3 (2)</td>
<td>1 (1)</td>
<td>2 (14)</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total PF N =</td>
<td>13</td>
<td>59</td>
<td>75</td>
<td>79</td>
<td>768</td>
<td>994</td>
</tr>
</tbody>
</table>

(A) Difference between Old Spanish and 17th vs. 19th-century 1s: p ≤ .0081; X² = 3.0625. No other 1s. differences significant.

A look at animacy in search of volition meaning is less revealing. If we propose that volition was an early meaning of PF (following Bishop 1973:89; see Section 1.1.2), we would expect that subjects incapable of desire, i.e., inanimate subjects, would occur less often with the PF in earlier data sets. In Table 4.31, we see that singular inanimate subjects, incompatible with volition and intention meanings, have in fact shown, if anything, a slight decline since Old Spanish. These results suggest that volition had no role in early PF development. Taken in tandem with the results for first-person uses, these results support the conclusion that intention, and not volition, was an important semantic bridge for early PF development (Bybee, Pagliuca & Perkins 1991), a meaning whose salience peaked in the 19th century.
4.4.4 Locatives

Another contextual clue to the generalization of the PF may be found in the PF’s rate of co-occurrence with locatives, as in (4.120). If the PF retains a motion meaning, it would be more likely to occur with locatives; as this meaning is lost, the rate of co-occurring locatives should decline (Pérez-Saldanya 2003:72; cf. Torres Cacoullos 2000:Ch. 3).

(4.120) Que tarde o temprano se **va a venir** a España, él sueña con España… (COREC, CBCON014A, 20s)
'Which sooner or later he's going to come (PF) to Spain, he dreams of Spain'

Along with locative adverbials that occur clause-internally, other elements in the surrounding text or discourse may explicitly mention *place*, as in (4.121). Here, the location (**abajo en las ordenadores**) does not occur as an adverbial within the same clause and as such cannot be counted among the clause-internal locatives, yet it clearly designates the spatial area in which the action is to take place. These occurrences were noted separately, as clause-external locatives.

(4.121) **Voy a buscar** a mi hermano que está **abajo en los ordenadores** (COREC, CCCON005A, 20s)
'I'm going to get my brother, who is downstairs on the computers'

Table 4.32 depicts the rate of clause-internal locative adverbials and of clause-external locative elements by century.
Table 4.32. Distribution of locatives with all PF by data set

<table>
<thead>
<tr>
<th></th>
<th>Internal</th>
<th></th>
<th>External</th>
<th></th>
<th>Total locatives</th>
<th></th>
<th>Total PF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>OldSp</td>
<td>2</td>
<td>12</td>
<td>0</td>
<td>--</td>
<td>2</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>19</td>
<td>32</td>
<td>9</td>
<td>15</td>
<td>28</td>
<td>47</td>
<td>59</td>
</tr>
<tr>
<td>19</td>
<td>7</td>
<td>9</td>
<td>3</td>
<td>4</td>
<td>10</td>
<td>13</td>
<td>77</td>
</tr>
<tr>
<td>20w</td>
<td>10</td>
<td>12</td>
<td>3</td>
<td>4</td>
<td>13</td>
<td>16</td>
<td>80</td>
</tr>
<tr>
<td>20s</td>
<td>125</td>
<td>16</td>
<td>6</td>
<td>&lt;1</td>
<td>131</td>
<td>16</td>
<td>775</td>
</tr>
<tr>
<td>N=</td>
<td>163</td>
<td>21</td>
<td>184</td>
<td></td>
<td>1007</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.32 shows that the rate of co-occurring locatives reached its peak at 47% (28/59) in the 17th century, the earliest data set with an appreciable number of tokens. In the 19th century and thereafter, locative rates are low, between 13% in the 19th century and 16% in the 20th century in both written and spoken data sets (13/80 and 131/775, respectively).

These results reveal a decline in both clause-internal and clause-external locative co-occurrence between the 17th and 19th centuries. In the case of clause-internal locatives, there is a drop from 32% (19/59) to 9% (7/77), which is significant at the .01 level ($p \leq .0007$, $x^2 = 11.54008$); in the case of clause-external locatives, the rate drops from 15% (9/59) to 4% (3/77), significant at the .05 level ($p \leq .0206$, $x^2 = 5.356642$). These tendencies also meant a general drop in co-occurring locatives, from 47% (28/59) to 13% (10/77), which demonstrates a significant loss of allative meaning ($p < .00001$, $x^2 = 19.71378$). The occurrence of the PF with clause-internal locatives in 20th-century speech, at 16% (125/775), is not evidence for retention of motion meaning: the SF occurs with locatives at the same rate of 16% (62/379) in the same data set.\(^{82}\)

---

\(^{82}\) Of the SF occurrences with locatives, 22% (N=14) were a, 3% (N=2) were de, 26% (N=16) were en, 8% (N=5) were 'here', 16% (N=10) were por; 13% (N=8) were 'there', and 12% (N=7) were categorized as 'other'.

149
use of clause-external locatives with PF after the 17th century to less than 1% (6/775) in 20th-century speech also matches the SF’s 20th-century speech rate of less than 1% (2/379) co-occurring clause-external locatives.

Along with a significant decline in the general rate of locative adverbials with PF, we may posit that the kind of locatives with which it co-occurs have shifted as well. The data revealed six common types of locative adverbials that co-occurred clause-internally with PF. These included: indication of movement toward a goal in a ‘to’ or hacia ‘toward’ as in (4.122); indication of movement away from a point of origin in de ‘from’ or desde ‘from/since’ in (4.123); deictic reference not at the place of reference in allí ‘there’, ahi ‘there’ or allá ‘over there’ in (4.124); deictic reference to the point of reference in aquí ‘here’ in (4.125), reference to stationary location in en ‘in’ in (4.126); and reference to motion within a delimited space in por ‘around/by’ in (4.127). Other uses included other position adverbials, such as abajo ‘below’, as in (4.128).

(4.122) **vov a llegar** tarde a casa y el viejo, que esta medio chalao, va a armar la de San Quintin. (Billy, p. 30, 20w)
'I’m going to arrive late at home and the old guy, who is a bit nuts, is going to have a fit.'

(4.123) Desde la fiesta me **vov a tomar** la ceniza. (Conjuración, Act IV, Scene III, 19e.)
'From the party I’m going to take (PF) the ash.'

(4.124) **Allí te vov a aguardar.** (Príncipe, Act II, fol. 8r, line 410, 17c.)
'I’m going to wait (PF) for you there.'

(4.125) Pedro, soy Lulo, estoy en Barcelona, **vov a estar aquí** un mes. (COREC, CPCON006A, 20s)
'Pedro, it’s Lulo, I’m in Barcelona, I’m going to be (PF) here a month.'

(4.126) Ahora **vov a intentar** inscribirle en la Seguridad Social, para que me lo vigilen (Pobres diablos, p. 96, 20w)
'Today I’m going to try (PF) to register him in Social Security, so they’ll watch him for me'
(4.127) No, por el cuartel no lo **van a encontrar**. (Matar, p. 151, 20w)
'No, you aren't going to find (PF) him around the barracks.'

(4.128) digo pues me **voy a bajar** abajo y si no te localizo... (COREC, CCCON016A, 20s)
'I say well I'm going to go down (PF) below and if I don't find you...'

We see the rate of each of these locative types in Table 4.33.\(^{83}\)

**Table 4.33. Distribution of types of clause-internal locative in all PF by data set**

<table>
<thead>
<tr>
<th>Locative</th>
<th>17 % loc</th>
<th>% PF</th>
<th>19 % loc</th>
<th>% PF</th>
<th>20w % loc</th>
<th>% PF</th>
<th>20s % loc</th>
<th>% PF</th>
</tr>
</thead>
<tbody>
<tr>
<td>a, hacia</td>
<td>32</td>
<td>10</td>
<td>14</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>32</td>
<td>5</td>
</tr>
<tr>
<td>de, desde</td>
<td>--</td>
<td>--</td>
<td>14</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>&lt;1</td>
</tr>
<tr>
<td>there</td>
<td>16</td>
<td>5</td>
<td>14</td>
<td>1</td>
<td>20</td>
<td>2</td>
<td>4</td>
<td>&lt;1</td>
</tr>
<tr>
<td>en</td>
<td>32</td>
<td>10</td>
<td>43</td>
<td>4</td>
<td>30</td>
<td>4</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>por</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>20</td>
<td>2</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>here</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>10</td>
<td>1</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Unknown</td>
<td>21</td>
<td>7</td>
<td>14</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>32</td>
<td>100</td>
<td>9</td>
<td>100</td>
<td>12</td>
<td>100</td>
<td>16</td>
</tr>
<tr>
<td>Total N</td>
<td>19</td>
<td>59</td>
<td>7</td>
<td>77</td>
<td>10</td>
<td>80</td>
<td>125</td>
<td>775</td>
</tr>
</tbody>
</table>

Of the locatives in Table 4.33, some are more compatible with the lexical allative meaning of *ir*, while others are not. For instance, with locatives in *a* 'to', *hacia* 'toward', *de* 'from', and *desde* 'from/since', a trajectory of motion is profiled; in the first two, an agent is moving toward a goal, in the latter two, away from an origin. In the case of locatives such as *alli* 'there' or *allá* 'over there', motion can be inferred since in most cases motion will be necessary to arrive at a deictic 'there' from deictic 'here'. On the other hand, with the locatives in *por* 'around/by' and *en* 'in', along with the deictic *aqui* 'here', no trajectory is profiled.

With this in mind, then, we might propose that the early PF would tend to occur more with locatives that profile a trajectory, since this is the main component of the PF's

\(^{83}\) Old Spanish data are not included in this discussion due to the low token count of co-occurring locatives with PF in Old Spanish (N=2).
original lexical meaning: "agent on a path toward a goal." We would further predict that
the PF would tend to co-occur less with locatives that denoted no particular trajectory,
and that these tendencies would become weaker over time. Table 4.34 shows the
proportion of locatives co-occurring with PF that profiled a trajectory and the proportion
of those that did not. Here we see that "trajectory" locatives have consistently accounted
for approximately half (43-50%) of all co-occurring locatives in the past seven centuries,
with no evidence of decline. Non-trajectory locatives occurred with the PF in these data
since the 17th century, with a rate of 32%. In the 19th and 20th centuries, this proportion
rises to a rate of 40-43%, which remains consistently and slightly lower than the
proportion of trajectory locatives.

Table 4.34. Association of clause-internal locatives with trajectory motion in PF by data
set

<table>
<thead>
<tr>
<th>Loc. semantics</th>
<th>17 % loc</th>
<th>19 % loc</th>
<th>20w % loc</th>
<th>20s % loc</th>
<th>17 % PF</th>
<th>19 % PF</th>
<th>20w % PF</th>
<th>20s % PF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trajectory</td>
<td>47</td>
<td>15</td>
<td>43</td>
<td>3</td>
<td>50</td>
<td>3</td>
<td>48</td>
<td>6</td>
</tr>
<tr>
<td>No trajectory</td>
<td>32</td>
<td>10</td>
<td>43</td>
<td>4</td>
<td>40</td>
<td>7</td>
<td>42</td>
<td>8</td>
</tr>
<tr>
<td>Unknown</td>
<td>21</td>
<td>7</td>
<td>14</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Total %</td>
<td>32</td>
<td>100</td>
<td>9</td>
<td>100</td>
<td>12</td>
<td>100</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Total N</td>
<td>19</td>
<td>59</td>
<td>7</td>
<td>77</td>
<td>10</td>
<td>80</td>
<td>125</td>
<td>775</td>
</tr>
</tbody>
</table>

Despite occurring with the same proportion of locatives as the 20th-century PF, the 20th-
century spoken SF displays a slightly different pattern, with 39% (24/62) occurring with
locatives that profile a trajectory, and 50% (31/62) that do not (with 11% having
unknown profiles). This shows that speakers prefer to use PF with locatives that profile a
trajectory—a preference consistent with its lexical origins—and the SF is preferred with
locatives that do not profile a trajectory, nice evidence for retention in grammaticization.
4.5 Temporal adverbials

The rate of occurrence of temporal adverbials with PF is yet another contextual clue into the early meanings of PF. If the PF always had future meaning, then we would expect that it would co-occur, even in its incipient stages in Old Spanish and the 17th century, with temporal adverbials at least some of the time. If, however, the alternative hypothesis is true, and future meaning was gained, or became more salient, as the construction grammaticized, we would expect to find an increase in the rate of co-occurring temporal adverbials. Table 4.35 shows the rate of co-occurring specific and non-specific temporal adverbials with PF by data set. As we can see, the PF rarely co-occurred with temporal adverbials until the 19th-century Peninsular Spanish data, less than 1% of the time before the 19th-century data, and at rates ranging from 13-18% in the 19th- and 20th-century data sets.

Table 4.35. Rate of occurrence of temporal adverbials with all PF by data set

<table>
<thead>
<tr>
<th></th>
<th>Specific</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>Total PF</th>
</tr>
</thead>
<tbody>
<tr>
<td>OldSp</td>
<td>0</td>
<td>0</td>
<td>--</td>
<td>0</td>
<td>--</td>
<td>0</td>
<td>--</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>2</td>
<td>&lt;1</td>
<td>3</td>
<td>&lt;1</td>
<td>59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>10</td>
<td>5</td>
<td>13</td>
<td>14</td>
<td>18</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20w</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>10</td>
<td>13</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20s</td>
<td>68</td>
<td>9</td>
<td>8</td>
<td>133</td>
<td>17</td>
<td>775</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=</td>
<td>83</td>
<td>77</td>
<td>160</td>
<td>1007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is tempting to argue here that this increase shows unequivocally the acquisition of new meaning for the PF, perhaps through spatial > temporal metaphor (Sweetser 1988). I would like to offer, however, an alternative hypothesis. The rate of co-occurring temporal adverbials also increased with the SF during this time period (see Figure 6.9), though there is no reason to believe that the SF was undergoing metaphorical extension, though
it may have undergone "infusion of new meaning from the context" (Bybee, Perkins & Pagliuca 1994:6). It is nonetheless possible that competition within the same functional space could have led to an increase in the rates of temporal adverbials with both constructions as the default future form was being renegotiated (see Chapter 6). Poplack and Malvar argue that the lower rate of temporal adverbials with PF in Brazilian Portuguese is a reflection of its default status (forthcoming), an argument which is also appropriate here. This situation differs from that of locatives, then, since the path of PF grammaticization did not lead it into competition with other motion-based constructions.

4.6 Summary

This chapter has shown that, as the PF spread into new contexts, it became more capable of performing tasks regarding general prediction. The PF increased in type-token frequency (Section 4.4.2) and emerged into contexts that were semantically incompatible with its original allative meaning, such as with stative verbs (Section 4.4.1) and with locatives indicating no trajectory of motion (Section 4.4.4). During this development, the PF came to express intention, a meaning most salient in the 19th century, and which declined as the PF took on more general prediction uses; no association was found between the PF and volition (Section 4.4.3). Section 4.5 shows the general increase in rates of co-occurring temporal adverbials with PF, an increase I attribute to competition for default future status that occurred in the 20th century. Chapter 5 will describe the nature of the competition of the PF with the SF during its development, and we will see more clearly how the contextual generalization of PF into general futurity meant a gradual and growing encroachment upon SF territory.
5 Forms in competition

5.1 Language variation in use

So far, this study has looked at two future forms in Spanish, the Synthetic Future and the Periphrastic Future. We have seen that both of these differ in their lexical origins, chronological age, level of syntheticity, and semantic nuances, which follow from their lexical origins. These two forms, however, are not strangers, nor have they developed in isolation. Both forms share a key component: they are used as futures.

The variation found in Spanish future expression is a normal part of language, which is characterized by, as Labov (1972:204) has called it, "inherent variation":

"grammar is not fixed and absolute with a little variation sprinkled on top, but it is variable and probabilistic to its very core" (Bybee and Hopper 2001:19). This means that we should not expect to find, at any moment in time, a linguistic system that displays categorical form-function symmetry, i.e. one form for each function and one function for each form (D. Sankoff 1988a). Instead, it is also possible to find form-function asymmetry.84 That these features are present in language systems suggests that, though "the natural condition of language" may be "to preserve one form for one meaning, and one meaning for one form" (Bolinger 1977:x), dynamic processes in linguistic systems are a reflection of speakers' constant negotiation of form-meaning pairs. These systems,

84 Natural language is not the only conventionalized symbolic system that has this characteristic. In fact, the history of human writing systems shows similar tendencies: in a single system, one form may exist with multiple functions (polyphony), or various forms with the same function may co-exist (homophony), often for thousands of years, perhaps indefinitely (Powell forthcoming).
steeped in variation, make the analyst's job much more complex (cf. Haspelmath 2003:211).

This complexity is not insurmountable, since the patterns found in language are not random: the heterogeneity we find in every aspect of language is structured and probabilistic (Weinreich, Labov & Herzog 1968:99-100; cf. Bayley 2002:117, 122). In this structured variability we find the key to the nature of language change: evidence of processes whose origins stretch back hundreds or even thousands of years beyond living memory and processes that will not reach completion until long after our great-grandchildren have died.

5.1.1 Grammaticization and language change

One of the most robust universal phenomena observable in the formation of grammar is the process of grammaticization (e.g. Meillet 1912/1965:131; Lehmann 1982; Bybee & Dahl 1989; Hopper & Traugott 1993; Bybee, Perkins & Pagliuca 1994) (see Section 1.2.1). The cross-linguistic validity of the theory of grammaticization is found in evidence that there are, on the one hand, recurrent diachronic paths in which certain lexical constructions tend to come to express certain grammatical meanings (e.g. Bybee & Dahl 1989; Bybee, Perkins & Pagliuca 1994; on futures, see Bybee & Pagliuca 1987), and, on the other hand, that these diachronic paths are characterized by a highly constrained set of concurrent linguistic changes in form and function (e.g. Heine, Claudi & Hünnehmeyer 1991). Some of the most important changes we may see during grammaticization are:

i. formal reduction, often accompanied by phonological reduction (Hopper & Traugott 1993:130-150)
ii. increased fixedness within the construction (Hopper & Traugott 1993:123)

iii. generalization to new contexts, which tends to co-occur with:

   a. increased token and type frequency (Hopper & Traugott 1993:103)
   b. generalization of meaning (Hopper & Traugott 1993:96-100)
   c. generalization of grammatical function, i.e. the application to an ever-widening array of morphosyntactic functions (Bybee & Pagliuca 1985; Hopper & Traugott 1993:100)

It is important to keep in mind that not any of these phenomena on its own is a necessary and sufficient criterion for grammaticization. On the contrary, these are simply empirically verifiable signs of grammaticization, and it is thus important to observe more than one of these phenomena before asserting that a certain form or construction is undergoing grammaticization. 85

As we saw in Chapter 3, the SF has undergone both formal reduction and increased fixedness. In Chapter 4, I presented evidence for increased fixedness in the PF construction as well, and others have noted formal reduction of this form (Melis 2006:928n36). Furthermore, both constructions display semantics consistent with the universal diachronic paths for future grams proposed by Bybee, Pagliuca and Perkins (1991). In this chapter, I will use quantitative methods in order to examine two fundamental (and obviously intertwined) areas: variant semantics and variant choice, first, and second, the local manifestation of the universals paths of grammaticization these competing constructions are following.

85 For a more comprehensive list of the diachronic symptoms of grammaticization, see Company (2003:51-52).
Regarding semantics and variant choice, as noted in Chapter 2, I have drawn on the conclusions of previous studies of Romance SF and PF, and determined overt contextual features that may indicate certain semantics, such as intention or temporal specificity. The other essential matter here, of course, is the process of grammaticization itself. For instance, as forms generalize into new contexts, does this happen in all contexts at the same time (Kroch 1989), or do speakers use the innovative construction in some contexts before others (e.g. Poplack & Malvar forthcoming)? A key component of the development of the PF, and its subsequent competition with the SF, is the question of its generalization into erstwhile SF territory. What we are dealing with here is a case of layering, such that these constructions compete for the same semantic space (Hopper & Traugott 1993:123).\(^86\)

The relationship of the paths of grammaticization of two or more competing constructions is not entirely clear. It is possible, for example, that the SF’s weakened ability to express intention allowed the PF to invade upon its erstwhile territory, taking on the task of expression intention and, subsequently, prediction. It is also possible that the semantic generalization of the PF into intention and prediction, a predictable path given its original semantics (Bybee, Pagliuca & Perkins 1991), triggered further grammaticization of the SF, pushing it out of the realm of competition (into mood). Or, of

---

\(^{86}\) This is not meant to imply that conservative forms are just as semantically robust as their innovative competitors; in fact, this is not the case: "the retention hypothesis predicts that newer constructions make a richer semantic contribution than older ones" (Torres Cacoullas 2000:166). These statements simply mean that there is no time in which speakers were left without options for expressing the meanings they wanted to express, e.g. there was not a moment in which Spanish speakers did not have some way of expressing futurity, which would remain the case with or without a specific construction used for future.
course, both phenomena could be occurring simultaneously: as speakers perceive the SF as semantically weaker, they might use the PF more, which would, in turn, strengthen the presence of PF as a contender for the expression of intention and prediction, leading speakers to choose it ever more often. Diachronic data seem to indicate that this complex dynamic may have been working in both directions simultaneously, as we will see in Section 5.3.5.1.

Generalization of a form, as mentioned above, is generally twofold: this process tends to entail generalization of meaning, from more concrete to more abstract, and generalization of function, such that the construction is appropriate in an increasing variety of contexts (Hopper & Traugott 1993). Generalization of function also means increased text frequency: as the construction becomes increasingly useful and occurs in more contexts, its overall rate of use increases, often dramatically. Some have viewed the process of semantic generalization as consisting primarily of the loss of semantic (and other) content (e.g. Heine & Reh 1984:15), or, more precisely, the loss of components of meaning through a process of "semantic bleaching" (e.g. Bybee & Pagliuca 1985). Hopper and Traugott (1993) argue that, at least in the early stages of grammaticization—which are, indeed, what we are observing with the PF in the 17th and perhaps even 19th century—the semantic changes involved are better characterized as shift as opposed to loss. Hopper and Traugott (1993:88) describe English's go-future, in many ways parallel to the Spanish PF, as a prime example:

For example, with reference to the development of future go, Sweetser says: "we lose the sense of physical motion (together with all its likely background influences). We gain, however, a new meaning of future prediction or intention—together with its likely background inferences" (Sweetser 1988:392). In speaking of the subjectification of be going to, Langacker draws attention to the loss of objective locational reference
points that movement entails, and suggests that this loss is replaced by realignment to the speaker's temporal perspective (1990:23). In other words, one meaning is demoted, another promoted. Nonetheless, the "promotion" of a meaning is simply a natural consequence of the loss of a component of meaning: as a once-salient aspect of meaning fades away, speakers began to associate the before not-so-salient components of meaning with the construction, which, depending on frequency, may or not become the most salient. Empirical evidence presented in Chapters 3 and 4 showed that both the SF and the PF underwent semantic changes in which a latent semantic feature—present though not fully exploited even for centuries—came into focus. In the case of the SF, a weakening of the ability to express future temporality and intention led to an epistemic modal use; in the case of the PF, a decline in the co-occurrence of locatives accompanied the bleaching of motion meaning in the purposive motion *ir a* + INF construction (cf. Torres Cacoullos 2000), leaving purpose (i.e. intention) and posteriority (i.e. prediction), the former of which appears to have subsequently bleached as well. In neither case is there evidence of a "new" meaning, but rather a loss of constraints on the contexts in which certain meanings could occur through semantic bleaching and concurrent contextual generalization.

In the examination of such a dynamic process of semantic and morpho-syntactic change, what must pique the interest of an analyst considering natural data is the possibility of empirically tracing the coinciding contextual generalization of an innovative construction and a conservative construction when there is layering. In this chapter, I will exploit the structured variation between the SF and the PF in order to trace the history of these forms not as isolates, but as dynamic parts interacting in an ever-changing system. Furthermore, I will focus on the question that has haunted all who have
researched this subject (cf. Matte Bon 2005): if both SF and PF are (competing) futures, as they are commonly recognized to be (Melis 2006:929n37), how do speakers choose one form over the other? Indeed, speaker choices are the driving question in all variationist work (Labov 1976).

Here I will offer empirical evidence that the lexical origins of the PF have conditioned the way it has become a future, and that the origins of the SF helped to determine the outcome when the process of renewal began, as predicted by Bybee and Pagliuca's (1987) notion of semantic retention (which, indeed, was first mentioned in their discussion of future forms). These findings echo those of Poplack and Tagliamonte (2001) for the case of going to in English, as well as for the diachronic shifts in future variation patterns found by Poplack and Malvar (forthcoming) in Brazilian Portuguese. I will show that, while overall tendencies in the choice between PF and SF have demonstrated great stability over hundreds of years, the PF has slowly encroached upon SF territory within the future realm. This encroachment is due to semantic and functional generalization, which has meant that centuries-old constraints of PF-SF variation are gradually losing ground, while newer ones are emerging.

The case of Spanish SF-PF variation presented here is in fact similar in many ways to the synchronic portrait of Canadian French future variation presented by Poplack and Turpin (1999), who found that the SF was most likely to occur with the formal person vous 'you', with specific temporal adverbs, and, most strongly, in negative polarity. A fundamental difference between the French variety studied by Poplack and Turpin (1999) and the Peninsular Spanish variety under investigation here is that the SF is still quite frequent and productive in the latter, unlike its American counterparts (e.g.
Cartagena 1995-1996; cf. Aaron 2003), and unlike the spoken Canadian French SF, which retains little productivity (Poplack & Turpin 1999:157, 160). They conclude that the PF has taken over the functional space originally occupied by the nearly defunct (in spoken language) SF, and that the PF's presumed earlier functional space, including a meaning of temporal proximity, is being replaced by the Futurate Present (Poplack & Turpin 1999:161-162).

The goal here is to discover the conditioning of PF-SF variation in Spanish, and any diachronic changes in conditioning and what these can tell us about how speakers choose one form over the other, with particular attention paid to linguistic features meant to operationalize the semantic subtleties that have been associated with these constructions in both prescriptive and descriptive literature. A variationist approach to grammaticization is in this case ideal, since it will offer access to subtleties in distributional patterns that are invisible with other methodologies. As mentioned above, structured linguistic variability and language change are inextricably linked; in fact, the latter depends on the former, since without variability there can be no change (D. Sankoff 1988a), and the former continuously recreates the latter, forming a kind of mutual determination (Company 2003:25). One of the linguistic consequences of language change is "layering," which occurs when a form develops new meanings without immediately (or ever) replacing other forms within the same functional domain (Hopper 1991:22). This is clearly the case in the Spanish future: the PF developed a future

---

87 Poplack and Turpin (1999:160) remark that, "With the exception of a number of fixed uses, IF [i.e. SF], whose productive use does not exceed 6 percent, is more appropriately characterized as a negative polarity item and/or a marker of formality than a future marker in spoken French."
meaning while the SF continued to be used as a future, changing the face of the functional domain of futurity.

Though forms do compete within one functional domain, this is not necessarily a competition among equals. In the case of the Spanish future, the competing forms have distinct lexical origins. In studying this layering, then, we must remember that, as forms generalize and undergo divergence (Hopper 1991), older meanings of the form do not disappear, but rather co-exist with newer meanings (Lichtenberk 1991; Hopper & Traugott 1993:3, Bybee, Perkins & Pagliuca 1994:17; Schwenter 1994:81). This means that we should expect to find retention (Bybee & Pagliuca 1987) (also called persistence in Hopper 1991) of older meanings. Divergence may take the obvious form of polysemy, where one form has more than one meaning (Company 2003:13). This is the case, as we saw in Chapter 3, for the SF, which has taken on a non-temporal epistemic meaning without ceasing to function as a future in other contexts.

The consequences of retention, however, may be subtler. Instead of (or alongside) divergence, there may be distinctive patterns of variation with competing forms: synchronic syntactic or semantic constraints on the variation between the grammaticizing form and competing forms (Torres Cacoullos 1999; Poplack & Tagliamonte 2001:207-235), e.g., on the choice of PF or SF in the realm of futurity, will reflect the competing constructions' lexical origins and diachronic trajectories. Section 5.3 will examine this hypothesis at some length.

5.1.2 Form-function asymmetry in the Spanish future

The majority of the literature on the Spanish and other Romance futures has paid particular attention to the differences between these forms. Taking a synchronic
approach, most scholars have maintained that these forms appear in different (mutually exclusive) contexts or with different meanings. These arguments demonstrate a certain discomfort with the notion of form-function asymmetry, a discomfort which may have its roots in Bolinger's claim that "the natural condition of language" is one form for one function, a claim that challenged the generativist notion of functional equivalence of different surface structures and that acknowledged the role of discourse and usage in meaning (Bolinger 1977:x). Building on this idea, however, some scholars within the usage-based tradition have explicitly acknowledged the possibility—centrality, even—of semantic neutralization between these constructions in some contexts (e.g. Fleischman 1982 on Romance; Cartagena 1995-1996:94 on Spanish; Poplack & Turpin 1999 on Canadian French; Poplack & Malvar forthcoming on Brazilian Portuguese).

For the moment, I will focus on the hypotheses presented in both qualitative and quantitative studies that were the precursors to variationist studies of this topic. I limit this discussion to these non-variationist studies, first, because the statistical significance and relative importance of the various proposed semantic differences have yet to be simultaneously examined, and second, because the findings of variationist studies are intrinsically complex and generally do not lend themselves easily to summary. Findings of other variationist studies on Romance futures (Almeida & Díaz 1998; Poplack & Turpin 1999; Poplack & Malvar forthcoming) will be incorporated instead into my explanation of hypotheses and coding of data.

The literature on Romance futures is not lacking in explanations for PF-SF variation (see Section 1.1.2; also, e.g., MatteBon 2005; Poplack & Malvar
forthcoming). For instance, the PF has been associated with reference to an action whose preparatory stages have already been accomplished (Vet 1994), as an inchoative, as in (5.129), where the interlocutor has already done something that may lead the picture to break, or with prospective aspect, meaning that it refers to an action that is about to be carried out, either in intention or "objective reality" (Cartagena 1995-1996:80), as in (5.130), repeated from Section 4.3.1, in which the speaker refers to overt signs that the subject is going to go out.

(5.129) Cuidado! **Vas a romper** el - el cuadro. (COREC, CACON026B, 20s)
'Careful! *You're going to break* (PF) the – the picture'

(5.130) Tú **vas a salir**, según las señas? (El afán, Act II, Scene III, line 52, 19c.)
'You're **going to go out** (PF), according to the signs?'

This hypothesis is consistent with the origins of the PF, which denoted an agent on a path toward a goal (e.g. Bybee & Pagliuca 1987; Bybee, Pagliuca & Perkins 1991; Hopper & Traugott 1993:3; Bybee, Perkins & Pagliuca 1994:5).

The PF is also claimed to differ from the SF in that it is used to refer to events that are **proximate** or that are **relevant to the present** (Cartagena 1978:384; Fleischman 1982:87, 181; Melis 2006), as in (5.131)-(5.133).

(5.131) ¿Quiieren que les cuente qué es lo que **vamos a tener** hoy en el programa? (COREC, CBCON032A, 20s)
'Do you want me to tell you what **we're going to have** (PF) today on the show?'

(5.132) porque me he acordado que ya que venía aquí me **voy a ir** ahora mismo a - a esa bodega de ahí de Azcona a comprar unos vinos para un médico (COREC, CBCON014B, 20s)

---

88 This section is intended to offer a summary of only the most recurrent themes in the literature in which scholars have attempted to pinpoint semantic differences between PF and SF. For a fuller review of meanings that have been claimed to be associated with these constructions, please refer to Section 1.1.2.
'because I remembered as I was on my way here I'm going to go (PF) right now to
– to that grocery store there in Azcona to buy some wines for a doctor'

(5.133) Y **vamos a ir** el miércoles a ensayar con la orquesta. (COREC, CCCON018B,
20s)
'And we're going to go (PF) **Wednesday** to practice with the orchestra'

Others have claimed that the PF is used to refer to **distant events** that, however, will
happen within a **specified time frame** (e.g. Helland 1997:73-75), as in (5.134).

(5.134) Luego que Yamaha va a volver - porque y , las motos - italianas están apretando
muy fuerte. - pues ya está a la altura de las Hondas - y **va a empezar** el año
próximo también a competir oficialmente - en el mundial de 250, entonces
Yamaha debe espabilarse. (COREC, CEC0006B, 20s)
'Since Yamaha is going to return – because and, Italian motorcycles are pushing
really hard. –well now it’s at the level of Hondas – and they're going to start (PF)
to compete officially **next year** too in the World 250, so Yamaha had better stay
on its toes.'

However, no consensus has been reached on whether the Romance PF and SF really
carry different meanings or if they are interchangeable (Poplack & Turpin 1999). Poplack
and Turpin (1999:139; cf. D. Sankoff 1988a) summarize nicely the reasons underlying
such debates:

Where a difference in function (or lack of substitutability) between forms
is suspected, proponents of unique form-function relationships attempt,
first, to define the distinction, and then to link it to categorical contextual
cocurrences. Where this fails, as we have seen to be the case for future
temporal reference, the alternating forms themselves are said to be the sole
bearers of the proposed functional, semantic or pragmatic distinction. And
the nature of this distinction, in the absence of categorical surface
correlates, is strictly a matter of individual interpretation.

Thus, we cannot be content with any analysis of PF-SF distinctions without overt
empirical evidence for such distinctions in patterns found in natural data. One of the
difficulties with these analyses is that many of the distinctions proposed, such as the idea
that a speaker may simply **perceive** the action as having a preparatory state already
completed (Cartagena 1995-1996:80) or as relevant to the present (e.g. Berschim
1987:103-104; Fleischman 1982:95-97; Gagnon 1990; Melis 2006), rely on knowing what the speaker perceives or intends to express. However, speaker intention is most often unavailable to the analyst, and as such, if there are no overt clues about what the speaker wishes to communicate, then features said to be embodied in the forms themselves cannot serve as the basis of an analysis (Du Bois 1987:811-812; D. Sankoff 1988a; Poplack & Turpin 1999:145-146).

One important piece of evidence that semantic isomorphy is an inadequate characterization of language use is that natural patterns reveal that there is no such dichotomous separation of the PF and the SF. As we will see, there are cases in which there are no overt contextual clues about the meaning of these forms, and they overlap, i.e., they appear to occur in the same contexts, sometimes from the same speaker speaking about the same situation with the same subject in the same time of reference during the same turn. The fact that these forms overlap should not be cause for surprise. What we find in the Spanish future today is simply layering (Hopper 1991), a common consequence of language change: two forms sharing space within the same semantic domain. If we take a look at the functional spaces said to be associated with the PF, we see that the PF is not in complementary distribution with SF, i.e. a look at natural data reveals a variable picture. Though no examples of SF were found with such clearly prospective meaning as in (5.129), it is possible to find the SF used to refer to what the speaker appears to perceive as an imminent event, as in (5.135), where the speaker's
perception of imminence is shown unequivocally in the adverbial *de forma ya inminente*
'immensely'.

(5.135)Pues dentro de este proyecto, "Olimpia Dos Mil" se ha entendido que este puede
ser eso, un proyecto más y se nos van a esa Semana Blanca Canaria en el Pirineo
y *de forma ya inminente estarán* por allí estos expedicionarios que van a ser
bastantes e insisto, también la prensa se dará por allí una vuelta, a ver qué pasa.
(COREC, CCCON012B, 20s)
'Well in this project, "Olympia 2000," it has been understood that this can be just
that, one more project and they go away to that Canary White Week in the
Pyrenees and *immensely* those tourists, who are going to be a bunch, will be (SF)
over there, and I insist, the press will also pass by, we'll see what happens'

The SF can also be found to be used to refer to events that are *proximate* or that are
*relevant to the present*, as in (5.136)-(5.138).

(5.136) Lo que pasa que hoy *estará* muy lleno a lo mejor. (COREC, CACON006D, 20s)
'The thing is today it'll be (SF) really full maybe'

(5.137)Muy bien, Domingo, nos hemos pasado de tiempo, así es que *volveremos* dentro
de un momento - con "Lucky Strike" (COREC, CECON006B, 20s)
'Fine, Domingo, we're over time, so we'll return (SF) in a moment – with "Lucky
Strike" [cigarettes]'

(5.138)Pues que - vamos a ver. El martes, yo *dare* clase a estos niños pues seguramente
de seis a siete (COREC, CCCON029D, 20s)
'Well - let's see. *Tuesday, I'll give* (SF) class to these kids well for sure from six
to seven'

And, once again, the SF can also refer to *distant events* that will happen within a *specified
time frame*, as in (5.139).

---

89 In an elicitation experiment in Spain, in which the stimulus was: "Una madre, viendo a su hija jugar y
correr, dirá," 79.5% of participants responded with the PF, as in "¡Niña! ¡niña! ¡No corras tanto! ¡Te *vas a
cae!*", while only 3.6% responded with the SF, strongly suggesting that SF use in this context, while
possible, is indeed less likely (Berschin 1986:303).
(5.139) Se dice - que Honda el año próximo no hará motos oficiales en doscientos cincuenta, al cabo de dos días dicen: "No, que habrá – dos". (COREC,
CECON006B, 20s)
'They say – that Honda next year won't make (SF) official motorcycles in 250, but
after two days they say: "No, there'll be – two".'

It appears, then, that the meaning differences between PF and SF—if, in fact, there are
any—are not perceptible through qualitative analysis of natural data alone. Instead, one
tangible point of entry for the discovery of possible differences may be in the distinct
lexical origins of these two forms. As noted in Section 5.1.1, older meanings tend to be
retained even as forms develop newer meanings. While there is evidence for such
retention in the proportion and kind of co-occurring locatives with the PF (Section 4.4.4),
it is possible that the retention of older meanings can also be shown in the diachronic
trajectory of PF-SF variation.

If this is the case for the PF, then, we will see evidence of its original "agent on a
path toward a goal" meaning (Bybee, Pagliuca & Perkins 1991) in its tendencies to occur
in certain linguistic contexts. More specifically, the following hypotheses can be
proposed:

i. PF will be used more than SF with subjects that are capable of autonomous
   motion, i.e. animate subjects.

ii. PF will be favored over SF when referring to a specified time, given the
    specificity implied in the notion of "goal."

iii. PF will occur less often than SF with verb classes that are inconsistent with the
    original allative meaning (e.g. stative, psychological) and the redundant ir (e.g.

169
Diachronically, as the PF distances itself from its original lexical meaning, we would expect to find a weakening of the effect of these constraints over time.

5.2 Results: Frequency

Before entering into the question of the tendencies in PF-SF variation over time, it is important to step back for a moment to look at the bigger picture. While both forms have, indeed, been possible as futures since the 13th century, it is by no means the case that they have been on an even playing field. In fact, a quick glance at Table 5.36, which shows absolute and relative frequencies for both forms when used as futures by century, reveals that the PF has emerged from a fledgling construction with nearly imperceptible frequency relative to the SF in Old Spanish, at less than 1%, to a thriving, even preferred, construction in the 20th century, at 66% in Peninsular Spanish spoken data.⁹⁰

---

⁹⁰ Compare these 20th-century results to those of Ravazzolo (1998), who examined the frequency of PF, SF, and Present (P) used as futures in written 20th-century Peninsular Spanish. In a sample of formal writing, taken from March 11-15, 1998 editions of the newspaper El País (258,316 words), 80% (1123/1410) of these were SF, 9% (130/1410) were PF, and 11% (157/1410) were P. In the novel Dias Contados, by Juan Madrid, which contained a great number of dialogues that the author and other native speakers felt approximated spontaneous speech (60,598 words), 37% (236/640) were SF, 38% (242/640) were PF, and 25% (162/640) were P (cf. Matte Bon 2005). Like my data, these results suggest that the PF is rising in relative frequency in Peninsular Spanish.
Table 5.36. Absolute and relative frequencies of temporal PF and SF by century, raw and normalized per 10,000 words

<table>
<thead>
<tr>
<th>Century (word count)</th>
<th>PF</th>
<th>Norm per 10,000</th>
<th>SF</th>
<th>Norm</th>
<th>Total</th>
<th>PF:SF Ratio (%PF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old Spanish (−311,000)</td>
<td>13</td>
<td>0.4</td>
<td>1828</td>
<td>58.8</td>
<td>1841</td>
<td>1:141 (&lt;1%)</td>
</tr>
<tr>
<td>17-18th c. (−232,000)</td>
<td>59</td>
<td>2.5</td>
<td>1270</td>
<td>54.7</td>
<td>1329</td>
<td>1:22 (4%)</td>
</tr>
<tr>
<td>19th c. (−71,500)</td>
<td>75</td>
<td>10.5</td>
<td>432</td>
<td>60.4</td>
<td>507</td>
<td>1:7 (12%)</td>
</tr>
<tr>
<td>20th c., written (−79,000)</td>
<td>79</td>
<td>10.0</td>
<td>212</td>
<td>26.8</td>
<td>291</td>
<td>1:3 (27%)</td>
</tr>
<tr>
<td>20th c., spoken (−242,000)</td>
<td>768</td>
<td>31.7</td>
<td>379</td>
<td>15.7</td>
<td>1147</td>
<td>1:0.5 (66%)</td>
</tr>
</tbody>
</table>

As we see in Table 5.36 (which repeats Table 2.1 but excludes non-temporal uses), the PF's normalized absolute frequency increased from 0.4/10,000 in Old Spanish texts to 31.7/10,000 in 20th-century speech. The diachronic blossoming of the PF was accompanied by a 20th-century decline in the temporal SF, which, after centuries of occurring at a steady rate of 55-60 per 10,000 words, drops to only 26.8 in 20th-century texts, and to an even lower 15.7 in 20th-century speech. As the temporal SF declines in absolute frequency, we see a similarly astonishing rise in PF relative frequency, from less than 1% in Old Spanish, 4% in the 17th, up to 12% in the 19th, and then jumping to 27% and 66% in written and spoken 20th-century data, respectively. Thus, in 20th-century speech, we see a reversal: the PF is now the more frequent than the SF as a future form. These results point to a drastic change in the expression of the Spanish future; all evidence seems to indicate that, at least in terms of frequency alone, the PF has encroached upon, and indeed is taking over, SF territory.

\[91\] This number does not include epistemic SF, which was on the rise in the 20th century. If epistemic SF were included in Table 2.1 for the 20th century, the normalized SF frequency would be 29.9 (N=236) for the written data and 20.6 (N=499) for the spoken, which is still less than half of the token frequency found in earlier data sets (ranging from 54.7 to 60.4). The relative frequencies would be 25% PF for the written data and 61% PF in the spoken.
5.3 Variable rule analyses

While the token and relative frequency data show what appears to be a gradual change, they tell us little about the nature of this change. For instance, they do not tell us if the PF has replaced the SF in all contexts at the same time, or if certain contexts have historically favored the PF. While Kroch's (1989:206) Constant Rate Hypothesis predicts that new variants will spread into all contexts at the same rate, Poplack and Malvar (forthcoming) find that this prediction does not hold: "forms do appear first in restricted contexts. They may proceed to spread to others, as eventually happened with IR [i.e. PF], or they may remain entrenched in their restricted contexts" (Poplack & Malvar forthcoming, emphasis in original).

The history of the internal landscape of the ever-changing Spanish future is still relatively unknown. Given such sweeping changes in relative frequency, we might imagine that there has been a similar magnitude of change in the internal structure of PF-SF variation. We might see this change reflected in how speakers choose one form over the other. The results of the similar study by Poplack and Malvar (forthcoming), however, while not predicting wild internal instability, also do not necessarily predict the maintenance of constraint hierarchies when innovative constructions are emerging or when conservative constructions are reaching the end of their path. Instead, they show that such complications can "perturb" the binary system, such that constraints may lose significance or even reverse direction. Given these results, Poplack and Malvar (forthcoming) suggest that Kroch's (1989) is only applicable in "periods in which the repertoire of variants remains stable, i.e. periods which feature neither incipient nor moribund forms."

172
Despite the fact that the present study focuses on only two forms, and thus is not a multi-variant account like Poplack and Malvar's, we will in fact see similar results in PF-SF variation, most notably in the 20th century. The Peninsular Spanish future is characterized by both a tenacious maintenance of certain constraints, particularly those linked to the lexical origins of the PF, and an instability that I will argue in Chapter 6 is linked to the semantic shifts in the SF in the 20th century. Quantitative analyses of PF-SF variation over time help to discern these at times conflicting tendencies. In this study, I have used variable rule analysis (see Section 2.5) to take a synchronic snapshot of the four data sets in which the PF shows more than minimal presence, 17th, 19th and 20th-century literature and 20th-century conversation. Each snapshot, presented in Sections 5.3.1-5.3.4, reveals the probabilistic constraints on the choice between PF and SF at that moment in time. By comparing these synchronic portraits, in Section 4.3.4, we may pinpoint the ways in which the constraints on this choice have changed or remained the same over time.

Each subsection below presents a table with the results of one binomial up-and-down multivariate analysis. All analyses were conducted with the PF as the so-called application value, or basis of measurement, i.e., how likely the PF was to occur instead of the SF. A column on the left marked "Percent PF" provides the percentage (relative frequency) of PF occurrences in the particular listed context. A second column, marked "Probability," provides the statistical likelihood that the PF will occur in said context, taking into account all of the other independent factors considered in the analysis. As such, the higher the probability weight, the more the PF is favored in this context. Probability weights of .50 indicate that both the PF and the SF are equally likely to occur.
Some probability weights are in square brackets, indicating that these factors were not statistically significant. The weights in brackets were taken from the first "stepping down" in the binomial up-and-down analysis (Poplack & Malvar forthcoming:n22), and are shown only for comparison of tendencies in direction of effect. The third column gives the total number of tokens—both PF and SF—that occurred in this context in the data, and the fourth shows the proportion of the data represented by this linguistic context.

5.3.1 The 17th century

The earliest data set that provided an PF occurrence rate high enough for statistical analyses of PF-SF variation was of the 17th-century. Table 5.37 shows the variable rule analysis (VRA) results for this data set. Here, the PF makes up only 4% of the data. Nevertheless, the occurrence of the PF even at this early stage in development is not random. Among the PF tokens used as futures during this time period, quantitative analysis reveals structured patterns of co-occurrence.92

---

92 Non-future uses of these forms, such as epistemic SF and habitual motion PF, were excluded from VRA analysis, since they did not fall inside the envelope of variation (see Section 1.4.4 for an explanation of exclusions from this analysis).
Table 5.37. *Variable rule analysis of the contribution of factors to occurrence of PF, 17th century*

Input probability: 0.031 (4%), N = 59/1329

<table>
<thead>
<tr>
<th>Factor Group</th>
<th>% PF</th>
<th>Probability</th>
<th>N</th>
<th>% data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verb class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motion (excl. <em>ir</em>)</td>
<td>9</td>
<td>.74</td>
<td>107</td>
<td>8</td>
</tr>
<tr>
<td>Dynamic (non-motion)</td>
<td>6</td>
<td>.62</td>
<td>599</td>
<td>46</td>
</tr>
<tr>
<td>Stative/percep./psych.</td>
<td>1</td>
<td>.33</td>
<td>593</td>
<td>46</td>
</tr>
<tr>
<td><em>Range</em></td>
<td></td>
<td></td>
<td>41</td>
<td></td>
</tr>
<tr>
<td><strong>Temporal specificity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>5</td>
<td>.56</td>
<td>1097</td>
<td>82</td>
</tr>
<tr>
<td>Present</td>
<td>1</td>
<td>.22</td>
<td>232</td>
<td>18</td>
</tr>
<tr>
<td><em>Range</em></td>
<td></td>
<td></td>
<td>34</td>
<td></td>
</tr>
<tr>
<td><strong>Polarity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affirmative</td>
<td>5</td>
<td>.55</td>
<td>1083</td>
<td>86</td>
</tr>
<tr>
<td>Negative</td>
<td>1</td>
<td>.24</td>
<td>189</td>
<td>14</td>
</tr>
<tr>
<td><em>Range</em></td>
<td></td>
<td></td>
<td>31</td>
<td></td>
</tr>
<tr>
<td><strong>Sentence type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interrogative</td>
<td>6</td>
<td>[.62]*</td>
<td>94</td>
<td>7</td>
</tr>
<tr>
<td>Declarative</td>
<td>4</td>
<td>[.49]</td>
<td>1235</td>
<td>93</td>
</tr>
<tr>
<td><strong>Clause type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subordinate</td>
<td>5</td>
<td>[.55]</td>
<td>448</td>
<td>34</td>
</tr>
<tr>
<td>Main</td>
<td>4</td>
<td>[.47]</td>
<td>881</td>
<td>66</td>
</tr>
</tbody>
</table>

Log likelihood = -223.076; *p* = .027; Chi-square/cell = 0.5203

*Square brackets [ ] indicate that this effect does not achieve statistical significance (see Section 1.4.5).

In the 17th century, three factor groups showed a statistically significant effect on nascent PF-SF variation. The most powerful of these—indicated by the highest range of 40—is that of verb class. In Table 5.37, we see that the most favorable verb class for early PF was motion verbs (excluding *ir*, which never occurred in PF in this data set), followed by other dynamic verbs, such as *buscar* 'look for' and *decir* 'say'. The least favorable context in terms of verb class for PF is that of stative, psychological and perceptual verbs, as in (5.140), which favor the SF, as in (5.141).

(5.140) Camina, pues, amigo Sancho, y **vamos a tener** en nuestra tierra el año del noviciado, con cuyo encerramiento cobaremos virtud nueva para volver al nunca de mí olvidado ejercicio de las armas. (Quixote, Cap. LXVI, fol. 254v, 17c.)

175
'Well, walk, Sancho friend, and we are going to have (PF) in our land the year of initiation, with whose closing we will gain new virtue to return to the for me unforgettable exercise of arms.'

(5.141) Pues tu casa en paz tendrás. (Dama boba, Act III, Scene XV, 17c.) 'Well you will have (SF) your house in peace.'

Only 1% of these verbs occurred with the PF, compared to 9% of motion verbs. This should come as no surprise, since the former are wholly inconsistent with the meaning of "agent on a path toward a goal" (Bybee & Pagliuca 1987; Heine 1993:60; cf. Melis 2006:941).

The most telling result here is that motion verbs—which partially overlap semantically with ir 'go'—are more favorable to PF in this early stage than are other dynamic verbs. Company (2003:26) notes that "en el inicio de la gramaticalización existe redundancia o alta afinidad entre los contextos de uso y la forma y significado innovador." These results, then, show us the key to the very beginning of PF generalization into future contexts. This is precisely the context in which a "motion" meaning for ir is also not required, since the "motion" meaning is already provided by whichever motion verb co-occurs in this context. This semantic harmony of ir in construction with motion verbs is the same as that found in the early Spanish progressive construction as ir + ndo grammaticized (Torres Cacoullos 1999, 2000). This is true, for example, with embarcarse 'embark' in (5.142), which provides motion meaning independent of the presence of vamos.

(5.142) Señor, nosotros somos dos capitanes de infantería española; tenemos nuestras compañías en Napoles y vamos a embarcarnos en quatro galeras que dizen estan en Barcelona, con orden de passar a Sicilia. (Quixote, Cap. LX, fol. 234r, 17c.) 'Sir, we are two captains of Spanish infantry; we have our companies in Naples and we are going to embark (PF) on four ships, that they say are in Barcelona with orders to go on to Sicily.'
In a similar example of the motion verb *partir* 'part' in a future context, but with SF, in (5.143), we see that there is no less motion meaning than in (5.142), despite the absence of *ir* 'go'.

(5.143) *y, assí, con licencia del duque mi señor, yo me partire luego en busca desse dessalmado mancebo, y le hallaré y le desafiaré y le mataré* (Quixote, Cap. LII, fol. 199r, 17c.)

'and so, with the permission of the duke my sir, I will part (SF) afterwards in search of that inhuman man, and I will find him and I will challenge him and I will kill him'

This co-occurrence of *ir*-based PF with a motion verb, then, while not completely redundant as it would be in the occurrence of *voy a ir* 'I am going to go', provides enough redundancy that the hearer may be more likely to interpret other, secondary meanings for the conjugated *ir*. This case is similar to Torres Cacoullos' (2001:465) findings for the early tendency for nondirectional motion verb to occur with the *andar + ndo* Spanish progressive, yet unlike Poplack and Tagliamonte's (2001:338) finding that in earlier varieties of English motion verbs disfavored the use of *going to*. It is important to note that while the PF is most likely to occur in this semi-redundant context in its earliest stages, it does not allow for the fully redundant occurrence with *ir*, which would require an abrupt semantic change.

The other two factor groups that show a significant effect on PF-SF variation in the 17th-century data set are temporal specificity and polarity. In the first case, we see in Table 5.37 that the PF is favored (.56) in contexts where no temporal adverbial is present, as in (5.144), and strongly disfavored (.22) where an adverbial is present, as in (5.145).

(5.144) «Por mí», replicó don Quixote, «miente tu, Sancho, quanto quisiéres, que yo no te yere a la mano; pero mira lo que vas a dezir.» (Quixote, Cap. XXXI, fol. 119r, 17c.)

"'For me," answered Don Quixote, "you lie, Sancho, as much as you want, I won't hold your hand, but watch what you are going to say (PF)""
Quién duda que dirán que soy un loco. (Príncipe, Act II, fol. 1v, line 22, 17c.)
'Who doubts that they will say (SF) that I am a crazy man.'

(5.145) porque si entramos en cuenta, señores míos -y esto que voy a dezir agora, lo quisiera dezir hechos mis ojos fuentes, pero la consideracion de nuestra desgracia y los mares que hasta aquí han lluido, los tienen sin humor y secos como aristas, y, assí, lo dire sin lagrimas93 (Quixote, Cap. XXXIX, fol. 149v, 17c.)
'because if we realize, my men (and this that I'm going to say (PF) now, I would like to say it with my eyes like fountains, but the contemplation of our disgrace, and the seas that they have rained before now have left them without humor and dry like shores, and as such, I will say it without tears)'

en tanto que este mi criado me desarma, os lo dire sin faltar vn punto a la verdad del caso. (Quixote, Cap. LXV, fol. 251r, 17c.)
'as soon as my servant here disarms me I will tell (SF) you, without missing the truth of the case a bit.'

In fact, the occurrence of PF with temporal adverbials in the 17th-century data was so miniscule (only 3 tokens: 1 specific, 2 nonspecific), that an attempt to separately analyze specific and nonspecific adverbials was recognized at once to be a futile endeavor.94 Nevertheless, the fact that lack of temporal adverbial modification favored use of the PF from the beginning could be interpreted as foreshadowing its eventual rise to default status as a future temporal marker (see Poplack & Malvar forthcoming for a similar argument regarding the same pattern for Brazilian Portuguese PF).

---

93 Note here the alternation of PF and SF: "porque si entramos en cuenta, señores míos -y esto que voy a dezir agora, lo quisiera dezir hechos mis ojos fuentes, pero la consideracion de nuestra desgracia y los mares que hasta aquí han lluido, los tienen sin humor y secos como aristas, y, assí, lo dire sin lagrimas."

94 Of the SF tokens, 82% (1041/1270) occurred with no temporal adverbial, 8% (102/1270) with a specific temporal adverbial, and 10% (127/1270) with a nonspecific temporal adverbial. This shows that the SF was associated with nonspecific adverbials even before the PF came to represent a large sector of future expression, though the difference between the rates of nonspecific and specific temporal adverbials in 17th-century SF approaches but does not reach statistical significance ($p \leq .0833$, $x^2 = 2.999703$).
The third of the significant factors, with the lowest magnitude of effect on 17th-century PF-SF variation is that of polarity. Table 5.37 shows that the PF is strongly disfavored in negative polarity contexts like (5.146) in the 17th century, with a probability weight of .24. In fact, only 1% (N=2) of negative contexts occurred with PF in the 17th-century data set.

(5.146) Voyme y no voy a dormir, porque no se eche de ver que luego pienso volver (Principe Ynocente, Act III, fol. 9v, line 531, 17c.)
'I'm leaving and I'm not going to sleep (PF), so that it is not forgotten that I plan on returning later'

Negative polarity, then, favored the SF, as in (5.147).

(5.147) No intentaré en mi vida, secretario, por Apolo, otra caza como aquella, si me fuese a la vida necesario. (Principe Inocente, Act II, fol. 1v, line 1, 17c.)
'I will not attempt (SF) in my life, secretary, by Apollo, another hunt like that one, if my life depended on it.'

These results line up with Poplack and Turpin's (1999) finding that Canadian French SF is highly favored in negative polarity contexts. Also like these scholars, I am unable to explain the reasons behind this tendency. The 17th century data are the only data set in which positive polarity is shown to significantly favor the PF; by the 19th century (see Section 5.3.2), the favoring effect of positive polarity had already disappeared, to later find itself reversed in the 20th century (see Section 5.3.4).

5.3.2 The 19th century

In the early 19th century, the factors affecting PF-SF variation shift as the PF's relative frequency rises from 4% to 15%. As seen in Table 5.38, however, the most influential factor group remains the same: verb class has the highest magnitude of effect, with a range of 33. Here, dynamic verbs, as in (5.148), constitute the most favorable context
(5.148) Mis compañeros y yo no deseamos otra cosa sino que vuestra rubicunda celsitud nos dé una patente firmada y sellada según estilo, en la cual se exprese que nuestras obritas, las ya publicadas y las que vamos a publicar, de las cuales y de sus autores han dicho y dirán los envidiosos críticos tantas perreras, son elegantes, doctísimas, incomparables, y de aquí arriba lo que pareciese conveniente añadir en su elogio. (Derrota, n.p., 18c.)

'My colleagues and I want nothing except that your healthy honor may give us a signed patent sealed as is the style, in which it is expressed that our little works, those already published and those we are going to publish (PF), about which, and about whose authors, envious critics have said and will say (SF) so many stupidities, are elegant, very cultured, incomparable, and anything above this that may seem convenient to add in its elegy.'

Table 5.38. Variable rule analysis of the contribution of factors selected as significant to occurrence of PF, 19th century

<table>
<thead>
<tr>
<th>Factor Group</th>
<th>% PF</th>
<th>Probability</th>
<th>N</th>
<th>% data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verb class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic (non-motion)</td>
<td>22</td>
<td>.65</td>
<td>233</td>
<td>47</td>
</tr>
<tr>
<td>Motion (excl. <em>ir</em>)</td>
<td>16</td>
<td>.55</td>
<td>55</td>
<td>11</td>
</tr>
<tr>
<td>Stative/percep./psych.</td>
<td>6</td>
<td>.32</td>
<td>206</td>
<td>42</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clause type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subordinate</td>
<td>22</td>
<td>.63</td>
<td>88</td>
<td>17</td>
</tr>
<tr>
<td>Main</td>
<td>13</td>
<td>.47</td>
<td>418</td>
<td>83</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Temporal specificity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific</td>
<td>19</td>
<td>[.58]*</td>
<td>53</td>
<td>10</td>
</tr>
<tr>
<td>Absent</td>
<td>15</td>
<td>[.52]</td>
<td>392</td>
<td>77</td>
</tr>
<tr>
<td>Nonspecific</td>
<td>7</td>
<td>[.33]</td>
<td>62</td>
<td>12</td>
</tr>
<tr>
<td><strong>Polarity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affirmative</td>
<td>16</td>
<td>[.52]</td>
<td>449</td>
<td>89</td>
</tr>
<tr>
<td>Negative</td>
<td>8</td>
<td>[.36]</td>
<td>58</td>
<td>11</td>
</tr>
<tr>
<td><strong>Sentence type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interrogative</td>
<td>15</td>
<td>[.52]</td>
<td>82</td>
<td>17</td>
</tr>
<tr>
<td>Declarative</td>
<td>15</td>
<td>[.50]</td>
<td>425</td>
<td>84</td>
</tr>
</tbody>
</table>

Log likelihood = -199.360; p = .044; Chi-square/cell = 0.8395

*Square brackets [ ] indicate that this effect does not achieve statistical significance.
The only other factor group shown to be statistically significant for the 19th-century data is, somewhat surprisingly, clause type, albeit with a comparatively low range of 16. This result is even more perplexing when we look at the probability weights: subordinate clauses, like (5.149) favor the PF, with a probability weight of .63. Though this effect is found to be significant in these data only for the 19th century, the same general trend appears throughout the history of PF-SF variation, with 20th-century literature being the only exception (see Section 4.3.5).

(5.149) si no disfruta campestres felicidades, gozará de los favores con que el destino ya a honrarle. (El afán, Act III, Scene VI, line 272, 19c.)
'if you don't enjoy earthly joys, you will enjoy the favors with which destiny is going to honor you (PF)'

El primero... y el último que te pediré ya en mi vida. (Conjuración, Act II, Scene III, 19c.)
'The first... and the last that I will ask (SF) of you (already) in my life.'

The favoring of the PF in subordinate clauses is not what is expected, given that the PF is undoubtedly the more innovative of the two future forms, and recent work has suggested a correlation between innovative forms and appearance in main clauses (Bybee 2001).95

As an explanation, one may hypothesize that the PF is not favored in subordinate clauses in general, but rather in a specific construction in which subordinate clauses occur. A closer look at the subordinate clauses, however, reveals no such peculiarity. As

---

95 A separate multivariate analysis that took into account possible discourse uses in coding clause type (see note 96), like the results presented here, also revealed the clause type factor group to be significant in the 19th century data set. The results showed the same trend, with the PF favored in subordinate clauses (weight=.58, relative frequency of PF=19%), and less favored in main clauses (weight=.49, relative frequency of PF=13%). The significance of constraints and relative magnitudes of effect remained the same as in the analysis presented in the main text.
seen in Table 5.39, subordinate clauses occurred in five different kinds of construction in these data: relative clauses as in (5.150), object complement clauses as in (5.151), in *si-* clauses (i.e. 'if' clause) as in (5.152), and in *porque* 'because', *aunque* 'although', and *sino que* 'but rather' clauses, as in (5.153), and with *que* 'that' with an unexpressed head, as in (5.154).

(5.150) Mira que un casamiento como el que **vas a hacer**, muy pocas le consiguen. (El sí, Act II, Scene II, 18-19c.)
'Look, a wedding like the one that you're going to do (PF), very few get that.'

Pasado mañana, por último día de carnaval, celebra el Dux un festín magnífico, a que **asistirán** sus consejeros y muchos miembros del senado (Conjuración, Act I, Scene III, 19c.)
'Day after tomorrow, on the last day of carnival, the Dux will celebrate a magnificent feast, which his advisors and many members of the senate will attend (SF)....'

(5.151) ¡Estoy pensando que no tienes hijos... y que no **vas a comprender**me! (Conjuración, Act III, Scene III, 19c.)

---

* See Thompson (2002) for a problematization of the category of complement clauses. Though I am aware of the polemic nature of this category, I have chosen to adopt traditional categories here. The main motivation for this decision is diachronic consistency: Thompson (2002) identifies her characterization of modern-day complement clauses as part of an ongoing diachronic process. What has still to be addressed is when this process began for Spanish, and at what moment these clauses lost category membership. The same can be said of other constructions, such as *es que* 'it's that' and headless *que* 'that', whose status is ambiguous. Nevertheless, in order to ensure the accuracy of results presented here, separate multivariate analyses were done for the 19th and 20th centuries. The results of these separate analyses, which only considered relative clauses, *como* 'like/as' clauses, and clauses with *aunque* 'although' and *porque* 'because' to be subordinate, excluding all possibly ambiguous tokens from this factor group (such as headless *que*, *si-* clauses, and all complement clauses), will be presented in footnotes throughout this chapter as appropriate. As will be seen, recoding of this factor group had no effect on the significance of constraints or relative magnitudes of effect for any data set.
'I'm thinking that you don't have children... and that you're not going to understand (PF) me!'

¿crees tú que llegaré ese momento? (Conjuración, Act II, Scene III, 19c.)
'do you think that that moment will arrive (SF),'

(5.152) ¿Por qué?... Si él va a volver, y sabe ya que yo estoy muriéndome... No me dejará así, no... ¿Cómo había de tener corazón para eso? (Conjuración, Act V, Scene VI, 19c.)
'Why?... If he is going to come back (PF), and he already knows that I'm dying... He won't leave me like that, no... ¿How would he have the heart for that?'

sólo se oye el murmullo del viento en este canal solitario... ¡Si no vendrá!... ¡Si le habrá sucedido alguna desgracia!... ¡No, Dios mío, no' (Conjuración, Act II, Scene II, 19c.)
'only the murmur of the wind is heard in this solitary canal... If he won't come (SF)! If something bad may have happened to him!... No, my God, no'

(5.153) pero [el autor] no debe olvidar... que su obra no va a leerse descansadamente, al amor de la lumbre, para pasar las largas noches de invierno; sino que va a representarse en el teatro, en que todo aparece desmayado y frío, si no hay acción, movimiento, vida. (Conjuración, Apuntes sobre el drama Histórico, 19c.)
'but [the author] should not forget...that his or her work is not going to be read in a leisurely way, in the love of the light, to pass the long winter nights; but rather that it is going to be presented (PF) in the theater, in which everything appears fainted and cold, if there is no action, movement, life.'

MARIDO.- Porque yo te la pondré en la frente. (Conjuración, Act IV, Scene III, 19c.)
'HUSBAND: Because I will put (SF) it on your forehead.'

(5.154) RUGIERO.- (Levantándose.) Que vais a firmar mi sentencia. (Conjuración, Act V, Scene IX, 19c.)
'RUGERIO – (Getting up.) That you all are going to sign (PF) my sentence.'

Déjela usted, señora; que ella responderá. (El sí, Act II, Scene V, 18-19c.)
'Leave her alone, ma'am; (that) she'll answer (SF).'

Despite low token frequencies, the distribution of these constructions across the PF and the SF, as shown in Table 5.39, is strikingly similar: the most frequent are object complement clauses and relative clauses, averaging at 39% (34/88) and 38% (33/88), respectively. Object complement clauses account for 36% (25/69) of subordinate clauses.
in SF, and 47% (9/19) of these clauses in PF. Relative clauses constitute 38% (26/69) of
SF subordinate clauses and 37% (7/19) of PF. Other types of subordinate clauses make
up the remaining 23% (21/88) of the subordinate clauses in 19th-century Peninsular
Spanish.

Table 5.39. *Distribution of kinds of subordinate clauses in temporal SF and PF, 19th
century*

<table>
<thead>
<tr>
<th>Kind of subordinate clause</th>
<th>Temp. SF % (N)</th>
<th>Temp. PF % (N)</th>
<th>Total % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complement</td>
<td>36 (25)</td>
<td>47 (9)</td>
<td>39 (34)</td>
</tr>
<tr>
<td>Relative</td>
<td>38 (26)</td>
<td>37 (7)</td>
<td>38 (33)</td>
</tr>
<tr>
<td><em>Porque/aunque/sino que</em></td>
<td>9 (6)</td>
<td>5 (1)</td>
<td>8 (7)</td>
</tr>
<tr>
<td>Headless <em>que</em></td>
<td>14 (10)</td>
<td>5 (1)</td>
<td>12 (11)</td>
</tr>
<tr>
<td>Si-clause</td>
<td>3 (2)</td>
<td>5 (1)</td>
<td>3 (3)</td>
</tr>
</tbody>
</table>

Total % = 100  
Total N = 69 19 88

It appears, then, that the syntactic characteristics of subordinate clauses in PF and SF
reveal little difference between these constructions.

Let us take a closer look, now, at the meanings expressed in the object
complement clauses, which often occur with heads that give the speaker's subjective
perspective of the event (Thompson 2002). These tokens, then, may offer a window into
speakers' construals of events as, for example, certain or uncertain. In Table 5.40, I have
listed the heads of the object complement clauses that occurred in both SF and PF in
these 19th-century Peninsular data, divided by semantics into the categories of
uncertainty, belief, certainty, and perception and saying. The heads that expressed
uncertainty that occurred in these data were *esperar* 'hope', *no saber* 'not know', *parecer*
'seem', *quién sabe* 'who knows', and *suponer* 'suppose'. The heads that expressed belief
included *apostar* 'bet', *creer* 'believe', *discurrir* 'deduce', *pensar* 'think', and *sospechar*
'suspect'. The heads that indicated certainty included asegurar 'assure', estar cierto 'be sure', no hay duda 'there is no doubt', olvidar 'forget', saber 'know', and ser claro 'be clear'.97 The verbs of perception and saying included advertir 'warn', anunciar 'announce', decir 'say', jurar 'swear', mirar 'look', persuadir 'persuade', prometer 'promise', and ver 'see'.98

As we see, the PF does not occur with any heads expressing uncertainty. Instead, it most commonly occurs with heads expressing certainty, accounting for 44% (4/9) of the data, and belief, making up 33% (3/9). The SF, on the other hand, is used in complementary fashion, most often, first, with heads that are verbs of saying or perception, at 40% (10/25), followed by heads expressing uncertainty, at 28% (7/25). These tendencies, then, suggest an association of PF with certainty and SF with uncertainty (see, e.g., Bishop 1973:89; Imbs 1968; Tlaskal 1978:206-207; Confais 1995; Almeida & Diaz 1998; Jensen 2002). Furthermore, the "promise" meaning cited by some scholars and grammarians for SF (cf. Matte Bon 2005), or perhaps intention (e.g. Bybee & Pagliuca 1987; Bybee, Pagliuca & Perkins 1991; Villa Crésap 1997:64, 96), is seen in the SF's high rate of co-occurrence with heads that are verbs of saying, including, in fact, prometer 'promise' and jurar 'swear'.

97 Olvidar 'forget' was included under "certainty" because forgetting something implies having accepted the truth value of what has been forgotten, thus making the speaker's construal of eventuality of the event a certain one.

98 I have included the total Ns for each head in Table 5.40 and Table 5.44 in the case of disagreement in the assignment of these categories, so that they may be recategorized as desired by others.
Table 5.40. Head verbs of object complement clauses with SF and PF, 19th century

<table>
<thead>
<tr>
<th>Head semantics</th>
<th>Temporal SF % (N)</th>
<th>Temporal PF % (N)</th>
<th>Total % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Uncertainty</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esperar 'hope'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>No saber 'not know'</td>
<td>2</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td>Parecer 'seem'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Quién sabe 'who knows'</td>
<td>2</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td>Suponer 'suppose'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td><strong>Belief</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apostar 'bet'</td>
<td>--</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Creer 'believe'</td>
<td>2</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td>Discurrir 'deduce'</td>
<td>--</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Pensar 'think'</td>
<td>--</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sospechar 'suspect'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td><strong>Certainty</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asegurar 'assure'</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Estar cierto 'be certain'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>No hay duda 'there's no doubt'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Olvidar 'forget'</td>
<td>--</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Saber 'know'</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Ser claro 'be clear'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td><strong>Perception/Saying</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advertir 'warn'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Anunciar 'announce'</td>
<td>--</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Decir 'say'</td>
<td>4</td>
<td>--</td>
<td>4</td>
</tr>
<tr>
<td>Jurar 'swear'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Mirar 'look'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Persuadir 'persuade'</td>
<td>--</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Prometer 'promise'</td>
<td>1</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Ver 'see'</td>
<td>2</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td><strong>Total N</strong></td>
<td>25</td>
<td>9</td>
<td>33</td>
</tr>
</tbody>
</table>

I will return to the question of certainty in complement clauses in my discussion of the results for 20th-century spoken data (where token frequency is higher and thus facilitates generalization), in Section 5.3.4, where I will present another view of PF-SF relative frequency according to head semantics.  

99 Note that all other factor groups in the 19th-century results, though not selected as significant, showed the same direction of effect as in the 17th-century results.
5.3.3 The 20th century, written data

The results for the 20th-century written data are shown in Table 5.41.\(^{100}\) Here, the PF constitutes over one-quarter of the data (27%). Following suit with previous time periods, the written 20th-century PF-SF variation is constrained most strongly by verb class. Both motion verbs and other dynamic verbs favor the PF in nearly equal measure, with probability weights of .60 and .59, respectively. Stative verbs, verbs of perception and psychological verbs continue to disfavor the PF, with a probability weight of .34.

Strikingly, \(\text{ir}\) is still absent as an option with PF, occurring instead categorically with SF.

In the 20th century, adverbials come into play once again, much in the same way they did in the 17th century, with the PF favored when no adverbial is present, as in (5.155), with a probability of .56.

(5.155) \(\text{Está tan flaca que da pena verla. Los vecinos }\text{van a pensar}\text{ que la matamos.}\)
\(\text{(Billy, p. 5, 20w)}\)
\('\text{She's so thin that it's pitiful to see her. The neighbors are going to think (PF) that we're killing her.'}\)

Furthermore, while both specific and nonspecific temporal adverbials disfavor the PF, the PF is more likely to occur with specific temporal adverbials, as in (5.156), at a probability of .38, compared to .31 for nonspecific temporal adverbials, as in (5.157).

(5.156) \(\text{Esta vez }\text{voy a convertir el pellejo de alguno de ellos en una regadera. (Billy, p. 36, 20w)}\)

\(^{100}\) A separate multivariate analysis that took into account possible discourse uses in coding clause type (see note 96), like the results presented here, revealed the clause factor group not to be significant. It should be noted further that preliminary results showed the same trend, with the PF less favored in subordinate clauses (weight=[.29], relative frequency of PF=12%), and more favored in main clauses (weight = [.52], relative frequency of PF=28%). No other constraint hierarchies or relative magnitudes of effect differed from the analysis presented in the main text.
This time I'm going to turn (PF) the skin of one of them into a showerhead.'

(5.157) Que no salgas con mi hermano, que un día te vas a volver tísico con esos maratones que te pegas. (Matar, p. 143, 20w)
'Don't go out with my brother, because one day you're going to get (PF) sick with those marathons that you put yourself through.'

In contrast, the SF is favored most strongly, then, with nonspecific temporal adverbials, as in (5.158), followed by specific temporal adverbials, as in (5.159), and disfavored when there is no temporal adverbial present, as in (5.160).

(5.158) Algún día tendrá que rendirse a la evidencia... (Billy, p. 6, 20w)
'Someday he will have (SF) to give in to the evidence'

(5.159) No arrancarán antes de las doce. Hay siempre alguien que llega tarde. (Billy, p. 49, 20w)
'They won't get started (SF) before twelve. There is always someone who arrives late.'

(5.160) Vuestra sangre lavaré mi honor y dejaré de ser la risa de los empleados de mi garaje. (Billy, p. 35, 20w)
'Your blood will wash (SF) my honor and I will no longer be the laughing stock of the employees of my garage.'

It is curious to see that this constraint remains so strong; its relative magnitude of effect has increased, with a range of 25, making it the second-most influential factor in PF-SF variation in the 20th-century Peninsular Spanish written data. It is important to remember, however, that the space in which PF and SF compete is not the entire picture: phenomena outside the temporal space of futurity may have had a hand in the shape of the variation in futurity. Chapter 6 will address these issues in depth, but suffice it to say that the strengthening of the effect of the association of PF with lack of temporal adverbial modification may be best understood as brought on by shifts in the semantics of the SF into use in non-temporal realms and an intensification of the favoring of the
temporal SF in contexts with a nonspecific temporal adverbial as it lost default status.\textsuperscript{101}

The most common adverbials that occurred with SF were 
\textit{ya} 'already' at 40\% (17/43), \textit{nunca} 'never' at 9\% (4/43), and \textit{cuando} 'when' clauses, also at 9\% (4/43).

Table 5.41. \textit{Variable rule analysis of the contribution of factors selected as significant to occurrence of PF, 20\textsuperscript{th}-century written data}

\begin{table}
\centering
\begin{tabular}{lllll}
\hline
Factor Group & \% PF & Probability & N & \% data \\
\hline
\textbf{Verb class} & & & & \\
Motion (excl. \textit{ir}) & 38 & .60 & 18 & 6 \\
Dynamic, non-motion & 33 & .59 & 170 & 59 \\
Stative/percep./psych. Range & 15 & .34 & 100 & 35 \\
\textbf{Temporal specificity} & & & & \\
Absent & 31 & .56 & 219 & 75 \\
Specific & 17 & .38 & 23 & 8 \\
Nonspecific Range & 12 & .31 & 49 & 17 \\
\textbf{Sentence type} & & & & \\
Interrogative & 53 & .73 & 26 & 9 \\
Declarative Range & 24 & .48 & 265 & 91 \\
\textbf{Polarity} & & & & \\
Affirmative & 29 & .52 & 246 & 84 \\
Negative & 16 & .38 & 45 & 16 \\
\textbf{Clause type} & & & & \\
Main & 28 & .50 & 229 & 79 \\
Subordinate & 24 & .48 & 62 & 21 \\
\hline
\end{tabular}
\end{table}

Log likelihood = -156.275; $p =$ .041; Chi-square/cell = 0.8991

*Square brackets [ ] indicate that this effect does not achieve statistical significance.

\textsuperscript{101} Absence of adverbial specification also favors of \textit{going to} vs. \textit{will} in Canadian English (Walker & Torres Cacoullos 2005) and PF (and later P) versus SF in Brazilian Portuguese (Poplack & Malvar forthcoming). My interpretation here is consonant with Poplack and Malvar's argument that this association is a reflection of the (eventual) default status of PF.
A shift comes with the tendency of the PF to occur with interrogatives showing statistical significance for the first time in this data set. As we see in Table 5.41, interrogatives, such as (5.161), favor the PF, with a probability weight of .73.

(5.161) SEÑOR: ¿Va a grabar las contestaciones?
JOVEN PERIODISTA: Pues sí. (Billy, p. 57, 20w)

'MAN: Are you going to record (PF) the answers?
YOUNG JOURNALIST: Well yes.'

This factor group shows the same magnitude of effect as temporal specificity, with a range of 25. While the significance of this factor group could, of course, be simply a reflection of the increased frequency of the PF in the 20th century, it also reflects an intensification of earlier tendencies. If not due simply to the number of tokens in this data set, this could be the result either of a rise in interrogative use with the PF or a decrease in interrogative use with the SF.

First, then, we may ask: is the increased magnitude of effect of the sentence type factor group the result of innovative PF semantics, or of a special meaning for PF? I would argue that the answer is both "yes" and "no." As noted in Chapter 4, the PF does often occur in modern usage in a special construction of incredulity, with a strong epistemic modality; this construction occurs most often in interrogatives (Melis 2006:921), as in (5.162), where the temporal reading of the PF is hardly salient (yet still possible).

(5.162) ¿Tú sabes acaso lo que pasa en La bruja animosa...? ¿No...? Pues, entonces...
Claro, qué vas a saber tú. (Pobres diablos, p. 37, 20w)
'Do you happen to know what's happening in La bruja animosa...? No...? Well, then... Of course, what are you going to know (PF).'

Of all tokens of interrogative PF uses in incredulity contexts, (5.162) is the only one in this data set whose future temporality is questionable. This is important, since examples
such as (5.162) are often cited as prototypical examples of this construction with PF (cf. Melis 2006). If we are more generous, and this use is defined in the broadest sense, as any use that occurs within a construction implying doubt or impossibility, to include examples like (5.163), in which the speaker expresses strong doubt through *qué mierda* 'what the hell [lit. what shit]', then this construction makes up 50% (7/14) of PF interrogative occurrences in this data set.

(5.163) *¿Qué mierda* de reconstitución *van a hacer* sin el fiambre? (Billy, p. 48, 20w)
*What the hell* kind of reconstitution are *they going to do* (PF) without the corpse?*

This figure may seem sufficient to explain this rise in PF interrogative frequency. Along with a move of PF *into* (a specific kind of) interrogative context, this change can also be understood as a move of the SF *out of* interrogative future contexts. In Table 5.42, we can see this pattern.

**Table 5.42. Distribution of interrogatives in PF and temporal SF by data set**

<table>
<thead>
<tr>
<th>Century</th>
<th>PF</th>
<th>Temporal SF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>13</td>
<td>0/1</td>
<td>--</td>
</tr>
<tr>
<td>15</td>
<td>3/12</td>
<td>25</td>
</tr>
<tr>
<td>17</td>
<td>6/59</td>
<td>10</td>
</tr>
<tr>
<td>19</td>
<td>12/75</td>
<td>16</td>
</tr>
<tr>
<td>20w</td>
<td>14/79</td>
<td>18</td>
</tr>
<tr>
<td>20s</td>
<td>112/765</td>
<td>15</td>
</tr>
</tbody>
</table>

Here we see that the 20th-century SF, both in written and spoken data sets, shows the lowest rates of interrogatives in the history of the form, at 6% (12/212) and 4% (15/392), respectively. Though these changes are subtle and hardly impressive, they are statistically significant (see Section 3.4.1). Since this is the crucial moment in which epistemic modal (non-future) SF uses are on the rise (see Section 3.4.2), it is possible that the move of interrogative contexts out of temporal SF is related to this semantic development. I will
delve into the question of such synchronized change, including this particular change, in Chapter 6.

5.3.4 The 20th century, spoken data

The results for the final data set are shown in Table 5.43. Once again, we see that verb class, temporal specificity, and sentence type are significant factors in predicting PF-SF variation. A fourth factor group, polarity, is also shown as significant for the first time since the 17th century. What is notably different is the ordering of the factor groups by magnitude of effect. Verb class, for the first time, has lost its place at the top; it is now the third most influential factor group. The hierarchy of constraints has remained the same, however, with motion and other dynamic verbs favoring the PF with weights of .56 and .58, respectively, and stative verbs, verbs of perception and psychological verbs continuing to disfavor the PF, with a weight of .36. We see, however, the appearance of *ir* with PF for the first time, seen in (5.164).102

(5.164) Venga, que **vamos a ir** al cine. (COREC, CACON006C, 20s)
'C'mon, 'cause we're going to go (PF) to the movies.'

---

102 Preliminary VARBRUL analyses with *ir* 'go' separated from other motion verbs revealed that, in 20th-century spoken data, *ir* showed very little effect, with a probability weight of .49. Due to its lack of direction of effect, it was subsequently included along with other motion verbs for this data set.
Table 5.43. *Variable rule analysis of the contribution of factors selected as significant to occurrence of PF, 20th-century spoken data*

Input probability: 0.692 (67%), N = 768/1147

<table>
<thead>
<tr>
<th>Factor Group</th>
<th>% PF</th>
<th>Probability</th>
<th>N</th>
<th>% data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temporal specificity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>73</td>
<td>.57</td>
<td>862</td>
<td>75</td>
</tr>
<tr>
<td>Specific</td>
<td>59</td>
<td>.42</td>
<td>112</td>
<td>9</td>
</tr>
<tr>
<td>Nonspecific</td>
<td>37</td>
<td>.22</td>
<td>173</td>
<td>15</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td><strong>Sentence type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interrogative</td>
<td>88</td>
<td>.78</td>
<td>133</td>
<td>12</td>
</tr>
<tr>
<td>Declarative</td>
<td>64</td>
<td>.46</td>
<td>364</td>
<td>88</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
<td></td>
<td>32</td>
</tr>
<tr>
<td><strong>Verb class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic (non-motion)</td>
<td>74</td>
<td>.58</td>
<td>599</td>
<td>52</td>
</tr>
<tr>
<td>Motion (incl. <em>ir</em>)</td>
<td>69</td>
<td>.53</td>
<td>155</td>
<td>14</td>
</tr>
<tr>
<td>Stative/percep./psych.</td>
<td>55</td>
<td>.36</td>
<td>393</td>
<td>34</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td><strong>Polarity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>74</td>
<td>.61</td>
<td>116</td>
<td>10</td>
</tr>
<tr>
<td>Affirmative</td>
<td>66</td>
<td>.49</td>
<td>1031</td>
<td>90</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td><strong>Clause type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subordinate</td>
<td>70</td>
<td>[.52]*</td>
<td>338</td>
<td>30</td>
</tr>
<tr>
<td>Main</td>
<td>65</td>
<td>[.49]</td>
<td>809</td>
<td>70</td>
</tr>
</tbody>
</table>

Log likelihood = -648.430; $p = .039$; Chi-square/cell = 1.1124

*Square brackets [ ] indicate that this effect does not achieve statistical significance.

If we look at temporal specificity—now the most influential factor in PF-SF variation (Range=35)—we again see the same pattern observed in the written 20th-century data: PF is favored when there is no adverbial (.57), and disfavored when there is a temporal adverbial present, slightly when there is a specific temporal adverbial (.42) and strongly when there is a nonspecific one (.22). As mentioned above, I would argue that these tendencies are a reflection of shifts in contexts of use with SF as the SF lost, and the PF gained, default status. Chapter 6 will look at the distribution of epistemic uses of SF in relation to the changes in 20th-century uses of SF as a future; we will see that the co-
occurrence of SF with nonspecific temporal adverbials more than doubled in the 20th century, a trend I hypothesize to be linked to the increase of SF use in non-temporal epistemic contexts and a loss of the association of the SF with future reference.

The second-most influential factor group in 20th-century spoken data, sentence type, is, I believe, also linked to changes in SF semantics outside the realm of the future. We see once again in Table 5.43 that interrogatives, as in the written data, favor the PF, with a probability weight of .78. Again, as in the 20th century written data (Section 5.3.3), there is a fairly strong presence of rhetorical incredulity uses of PF in interrogative contexts, as in (5.165), whose incredulous semantics are indicated by the speaker's use of cómo coño 'how in the hell' and (5.166), in which the speaker offers information as to why the statement esto es otoño 'this is autumn' is unbelievable, i.e. that it is very cold for autumn (realmente hace un frío espantoso 'it's really horribly cold'.

(5.165) Digo oye, que no sé si voy a poder ir, porque no encuentro las llaves; y claro, cómo coño me voy a ir de casa habiendo dejado las llaves del garaje, de la casa, de todo, de todo puestas por fuera el domingo, y ayer lunes, pues como he estado todo el día en casa (COREC, CACON006A, 20s) 'I say listen, that I don't know if I'm going to be able to go, because I can't find the keys; and clearly, how in the hell am I going to leave (PF) the house having left the keys to the garage, to the house, to everything, to everything, outside on Sunday, and yesterday Monday, well since I've been at home all day'.

(5.166) Pero cómo ya a ser esto otoño, si realmente hace un frío espantoso? (COREC, CBCON032A, 20s) 'But how can this be (PF) autumn, if it's really horribly cold?'

Rhetorical incredulity contexts make up 29% (34/118) of interrogative PF uses in this data set. If these 34 tokens were removed from this analysis, then PF would account for 85% (84/99) of interrogative occurrences in this data set, still significantly higher (p ≤ .0002, $x^2 = 13.49149$) than the average proportion of PF of 67% (768/1147). This shows that, despite the high proportion of PF interrogatives found in rhetorical incredulity
contexts, this use alone cannot account for the favoring of PF in interrogative contexts in this data set. It is unclear what the reasons behind the longstanding favoring of PF in interrogatives are. I will discuss a couple of possibilities in section 6.3.2; however, as we will see, neither intention meaning (associated with the PF, see section 5.3.5.1) nor uncertainty (not associated in general with the PF, see Table 5.44 below) alone, given parallel interrogative patterns for epistemic SF (associated with uncertainty but not intention, see Section 3.5), can reasonably explain this tendency.

The final factor selected as significant is polarity, with negative polarity favoring the PF with a probability weight of .61. The effect of this factor group is minimal (Range = 12), but it is interesting to note that this represents a reversal in the direction of effect seen in the other data sets, including the 17th century, where this factor group was also found to be significant. The reasons for this are not completely clear, though the rates of negative polarity with temporal SF do decrease significantly in the 20th century (see Section 3.5).

It will be noted that clause type has no statistical effect in this time period.103 Nevertheless, I would like to revisit the association with certainty main verb heads within subordinate clauses that seemed to be occurring in the 19th century (Section 5.3.2). Have these tendencies been lost as the PF has undergone contextual generalization? We have already seen (Chapter 3, Table 3.17) that temporal SF tends to occur in object

---

103 A separate multivariate analysis in which clause type was recoded to consider discourse factors (see note 96) revealed similar results: clause type was not selected as statistically significant, with the first stepping down revealing PF as slightly favored with subordinate clauses with a weight of [.53] (69% PF relative frequency) and less favored in main clauses with a weight of [.50] (66% PF relative frequency). No other differences in statistical significance of factor groups or relative magnitudes of effect were found.
complement clauses with heads that express certainty more often than epistemic SF.

Table 5.44 shows how the PF lines up in this matter. Here we see that only 17% (12/72) of PF occurrences in object complements in 20th-century speech occur with heads that imply uncertainty, such as no saber 'not know' or preguntarse 'wonder', compared to 49% (22/45) of temporal SF and 88% (15/17) of epistemic SF. In contrast, 32% (23/72) of PF tokens in this context occur with head verbs of belief, such as creer 'believe' or pensar 'think', compared to 11% (5/45) of temporal SF and 6% (1/17) of epistemic SF. The semantic context most strongly associated with the PF in object complement clauses is that of certainty, at 38% (27/72) of PF occurrences in the subordinate clause context, occurring with verbs such as ser 'be' and saber 'know'. The temporal SF also occurs a fair amount in certainty contexts, at 29% (13/45), while the epistemic SF occurs only 6% (1/17) of the time with head verbs indicating certainty.
Table 5.44. Heads of object complement clauses with temporal SF, epistemic SF and PF, 20th-century speech

<table>
<thead>
<tr>
<th>Head semantics</th>
<th>Temporal SF</th>
<th>Epistemic SF</th>
<th>Temporal PF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
<td>% (N)</td>
</tr>
<tr>
<td><strong>Uncertainty</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imagine ‘imagine’</td>
<td>2 (49)</td>
<td>3 (15)</td>
<td>--</td>
<td>5 (1)</td>
</tr>
<tr>
<td>No saber ‘not know’</td>
<td>14 (27)</td>
<td>12 (15)</td>
<td>8 (12)</td>
<td>34 (15)</td>
</tr>
<tr>
<td>Parecer ‘seem’</td>
<td>1 (1)</td>
<td>--</td>
<td>3 (1)</td>
<td>4 (1)</td>
</tr>
<tr>
<td>Preguntarse ‘wonder’</td>
<td>1 (1)</td>
<td>--</td>
<td>1 (1)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Suponer ‘suppose’</td>
<td>4 (27)</td>
<td>1 (1)</td>
<td>--</td>
<td>5 (27)</td>
</tr>
<tr>
<td><strong>Belief</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creer ‘believe’</td>
<td>3 (5)</td>
<td>1 (1)</td>
<td>19 (23)</td>
<td>23 (29)</td>
</tr>
<tr>
<td>Darse cuenta ‘realize’</td>
<td>--</td>
<td>--</td>
<td>1 (1)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Entender ‘understand’</td>
<td>--</td>
<td>--</td>
<td>1 (1)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Estar en ‘be under the impression’</td>
<td>1 (1)</td>
<td>--</td>
<td>--</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Pensar ‘think’</td>
<td>1 (1)</td>
<td>--</td>
<td>2 (3)</td>
<td>3 (3)</td>
</tr>
<tr>
<td><strong>Certainty</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asegurar ‘assure’</td>
<td>1 (13)</td>
<td>--</td>
<td>--</td>
<td>1 (13)</td>
</tr>
<tr>
<td>Constar ‘be evident’</td>
<td>--</td>
<td>--</td>
<td>1 (1)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Pasar ‘happen’</td>
<td>--</td>
<td>--</td>
<td>1 (1)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Resaltar ‘end up’</td>
<td>--</td>
<td>--</td>
<td>1 (1)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Saber ‘know’</td>
<td>--</td>
<td>--</td>
<td>9 (27)</td>
<td>9 (27)</td>
</tr>
<tr>
<td>Ser ‘be’</td>
<td>7 (29)</td>
<td>1 (1)</td>
<td>11 (19)</td>
<td>19 (19)</td>
</tr>
<tr>
<td>Ser claro ‘be clear’</td>
<td>2 (5)</td>
<td>--</td>
<td>--</td>
<td>2 (5)</td>
</tr>
<tr>
<td>Ser seguro ‘be sure’</td>
<td>2 (13)</td>
<td>--</td>
<td>1 (3)</td>
<td>3 (3)</td>
</tr>
<tr>
<td>Ser verdad ‘be true’</td>
<td>--</td>
<td>--</td>
<td>3 (3)</td>
<td>3 (3)</td>
</tr>
<tr>
<td>Tener en cuenta ‘keep in mind’</td>
<td>1 (1)</td>
<td>--</td>
<td>--</td>
<td>1 (1)</td>
</tr>
<tr>
<td><strong>Perception/Saying</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contar ‘tell’</td>
<td>--</td>
<td>--</td>
<td>2 (10)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Decir ‘say’</td>
<td>3 (29)</td>
<td>--</td>
<td>7 (10)</td>
<td>10 (10)</td>
</tr>
<tr>
<td>Mirar ‘look’</td>
<td>--</td>
<td>--</td>
<td>1 (1)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Ver ‘see’</td>
<td>2 (1)</td>
<td>--</td>
<td>--</td>
<td>2 (1)</td>
</tr>
</tbody>
</table>

| Total %                    | 100         | 100          | 100         | 100    |
| Total N                    | 45          | 17           | 72          | 134    |

Like the 19th-century data, these findings lend support to the studies that have claimed uncertainty meanings to be associated with the SF and certainty with the PF (e.g. Confaís 1995; Almeida & Díaz 1998; Jensen 2002; Villa-Crésap 1997), though these results also indicate that this tendency does not approach a "rule" by any means. Furthermore, the association of the SF with "promise" meaning found in the 19th century (Section 5.3.2) seems to have disappeared. Villa-Crésap (1997:55-64) found similar associations in his
quantitative study of 20th-century spoken New Mexican Spanish, for which, in his data as a whole, he reported a rate of 50% certainty for PF and 11% for SF, and a rate of 9% uncertainty for PF and 70% for SF. It appears that the occurrence of SF in non-future-reference epistemic contexts is contaminating future temporal reference tendencies for the SF by association, which is to be expected, since the same form is used in both future-reference and non-future-reference contexts.

5.3.5 The face of the two Futures: 300 years

Now that I have presented the results by century, it will be useful to look once again at these data compared across centuries. Table 5.45 shows the results of the variable rule analyses (VRA) for all four data sets. The first four columns, under "Relative frequency PF," show once again the same trend seen in Table 2.1 and Table 5.36: a sharp rise in PF frequency across the board. It is the second set of columns, under "Probability," on which I would now like to focus.
Table 5.45. Variable rule analyses of the contribution of factors selected as significant to occurrence of PF, 17th-20th centuries

<table>
<thead>
<tr>
<th>Factor Group</th>
<th>Relative frequency PF</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>17th</td>
<td>19th</td>
</tr>
<tr>
<td>Verb class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic (non-motion)</td>
<td>6</td>
<td>22</td>
</tr>
<tr>
<td>Motion</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>Stative/percep./psych. Range</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Temporal adverbial modification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Present</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Sentence type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interrogative</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Declarative</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polarity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affirmative</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>Negative</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clause type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subordinate</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>Main</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

17th c.: Log likelihood = -223.076; p = .027; Chi-square/cell = 0.5203
19th c.: Log likelihood = -199.360; p = .044; Chi-square/cell = 0.8395
20th w.: Log likelihood = -156.275; p = .041; Chi-square/cell = 0.8991
20th c. s.: Log likelihood = -648.430; p = .039; Chi-square/cell = 1.1124
*Square brackets [ ] indicate that this effect does not achieve statistical significance.

In exploring these results, it is important to distinguish two phenomena. First, there is retention of lexical PF meanings evidenced in PF-SF distribution patterns. However, since futurity is not the end of the road and these forms have not remained static, there is also innovation in distribution linked to semantic changes in the SF outside of the realm of future temporal expression. In this section, I should like to discuss the former in some detail. I also hope to distinguish these from the latter, which will be discussed in Chapter 6.
A striking characteristic of the past 300 years of PF-SF variation is its stability. Though the statistical significance of the factor groups varies by century, the stability in overall tendencies is remarkable. Drawing on the diachronic comparison of constraint hierarchies similar to the inventive technique introduced in Poplack and Tagliamonte (2001:5-8), Table 5.46 shows a visual representation of the direction of effect—regardless of statistical significance—of the contextual factors considered over time. Such comparison of the direction of effect over time portrays a more complete picture of variant selection during grammaticization (cf. Poplack & Tagliamonte 2001).

Table 5.46. Factors’ direction of effect by data set

<table>
<thead>
<tr>
<th></th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verb class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic (non-motion)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Motion</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Stative/percep./psych.</td>
<td>⊘</td>
<td>⊘</td>
<td>⊘</td>
<td>⊘</td>
</tr>
<tr>
<td><strong>Temporal specificity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No temporal adverbial</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Specific temporal adverbial</td>
<td>⊘</td>
<td>✓</td>
<td>⊘</td>
<td>⊘</td>
</tr>
<tr>
<td>Nonspecific temporal adverbial</td>
<td>⊘</td>
<td>⊘</td>
<td>⊘</td>
<td>⊘</td>
</tr>
<tr>
<td><strong>Sentence type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interrogative</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Declarative</td>
<td>⊘</td>
<td>⊘</td>
<td>⊘</td>
<td>⊘</td>
</tr>
<tr>
<td><strong>Polarity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affirmative</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>⊘</td>
</tr>
<tr>
<td>Negative</td>
<td>⊘</td>
<td>⊘</td>
<td>⊘</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Clause type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subordinate clause</td>
<td>✓</td>
<td>✓</td>
<td>⊘</td>
<td>✓</td>
</tr>
<tr>
<td>Main clause</td>
<td>⊘</td>
<td>⊘</td>
<td>⊘</td>
<td>⊘</td>
</tr>
</tbody>
</table>

✓ = tends to favor the PF more
⊘ = tends to favor the PF less
→ = shows no tendency in variation (Probability weight = .50)
Indeed, we see a similar pattern for temporal specificity in all four data sets, with the PF favored in contexts in which no temporal adverbial was present, and by specific more than non-specific temporal adverbials. A look at the other factors considered reveals a strikingly similar stability in constraints over time. Interrogatives always tend to favor use of the PF, while declaratives appear to neither favor nor disfavor this form. Clause type is significant only in the 17th century, but the PF tends to occur more in subordinate clauses than the SF in all periods except 20th-century literature, though the effect is minimal, showing a low range of only 16. Verb class also shows remarkable stability: dynamic verbs (e.g. *comer* 'eat', *buscar* 'look for') and motion verbs (e.g. *salir* 'go out', *entrar* 'enter') consistently favor the PF, while stative (e.g. *estar* 'be', *tener* 'have'), perceptual (e.g. *ver* 'see', *oír* 'hear') and psychological verbs (e.g. *creer* 'believe', *querer* 'want') consistently disfavor the PF. These patterns show evidence of the PF's original lexical meaning, which was inconsistent with the latter three verb classes.

The changes we see in PF-SF variation since the 17th century are subtler. There are two kinds of change worth noting here. First, there are changes in the statistical significance of given contextual factor groups, which obviously may be due simply to the quantity of data used in each data set, but which also may indicate dramatic changes in the relative frequency of a variant within a given context. Second, there are shifts in the relative magnitude of effect of the significant factor groups in each data set. If we take the 17th century to be a portrait of a nascent PF, we find a great deal of consistency with the semantics of the lexical origins of this construction. Favoring by dynamic verbs, positive polarity and lack of temporal adverbials, 17th-century PF was used to refer to dynamic, positive actions that were not explicitly placed in the future temporal
framework. All of these contextual features are entirely consistent with an "agent on a path toward a goal" meaning, and are highly predictable from this meaning. By the 19th century, as would be expected for a grammaticizing form, two of the three original constraints tied to the lexical origins of PF have been lost, with only verb class retaining its sway. Also, clause type is significant for the first and only time in these 19th-century data, with the innovative PF occurring more often in subordinate clauses than the SF.

The 20th century shows a reversal: constraints strengthen. The sentence type constraint, which had not reached significance in previous data sets, emerges as significant in the 20th century; furthermore, the adverbial constraint, which had lost significance in the 19th century, comes back stronger than ever. Poplack and Malvar interpret the same phenomenon in their Brazilian Portuguese data as evidence of the default status of PF (forthcoming); while I agree in essence with this interpretation, I believe that the best evidence for this is found outside the envelope of variation (see Section 6.4.3). The sentence type constraint, though always minimally present, also cannot be explained only in terms of semantic retention. Instead, we may also understand the strengthening of this constraint as directly connected to the shift in SF from future temporal contexts to non-future-reference epistemic contexts (see Chapter 3 and Section 6.4.2).

The second kind of change we see in PF-SF variation involves the relative magnitude of effect of each statistically significant factor. In the gradual process of grammaticization, we would expect that early constraints attributable to the lexical origins of the PF would weaken or disappear over time as the PF became a more generalized future. This does appear to be the case. Table 5.47 lists the factor groups by
century, in descending order of magnitude of effect. Factor groups not selected as significant for the respective dataset are shown in parentheses.

Table 5.47. Ordering of magnitude of effect by data set, in descending order

<table>
<thead>
<tr>
<th></th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb class</td>
<td>Verb class</td>
<td>Verb class</td>
<td>Adverbial</td>
<td></td>
</tr>
<tr>
<td>Adverbial</td>
<td>Clause type</td>
<td>Adverbial</td>
<td>Sentence type</td>
<td></td>
</tr>
<tr>
<td>Polarity</td>
<td>(Adverbial)</td>
<td>Sentence type</td>
<td>Verb class</td>
<td></td>
</tr>
<tr>
<td>(Sentence type)</td>
<td>(Polarity)</td>
<td>(Polarity)</td>
<td>Polarity</td>
<td></td>
</tr>
<tr>
<td>(Clause type)</td>
<td>(Sentence type)</td>
<td>(Clause type)</td>
<td>Clause type</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.47 shows that, of the three significant factors in the 17th century, only one—verb class—survived as a statistically significant constraint over the next 200 years. The verb class factor group shows a classic pattern for generalization, from a robust constraint in the 17th century to a relatively weak constraint in the 20th-century spoken data. The classic nature of the weakening of verb class constraints as the PF generalized is even clearer when we examine the distributional patterns of the PF within particular verb classes, shown in Table 5.48.

Table 5.48. Distribution of verb classes in PF and temporal SF by data set

<table>
<thead>
<tr>
<th>Verb class</th>
<th>13 PF</th>
<th>15 SF</th>
<th>17 SF</th>
<th>19 PF</th>
<th>20w SF</th>
<th>20s SF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic (non-motion)</td>
<td>100</td>
<td>53</td>
<td>67</td>
<td>.48</td>
<td>.64</td>
<td>.44</td>
</tr>
<tr>
<td>Static</td>
<td>--</td>
<td>27</td>
<td>--</td>
<td>30</td>
<td>5</td>
<td>34</td>
</tr>
<tr>
<td>Motion</td>
<td>--</td>
<td>7</td>
<td>17</td>
<td>6</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>Ir</td>
<td>--</td>
<td>3</td>
<td>--</td>
<td>3</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td>Perception</td>
<td>--</td>
<td>4</td>
<td>17</td>
<td>6</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Psychological</td>
<td>--</td>
<td>6</td>
<td>--</td>
<td>7</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>&lt;1</td>
<td>99</td>
<td>2</td>
<td>98</td>
<td>4</td>
<td>96</td>
</tr>
</tbody>
</table>

In Table 5.48, we see a gradual encroachment of the PF into the territories traditionally claimed by SF, which is to be expected in grammaticization (García 1985a, 1991; Kroch 1989; Company 2001, 2003). The most obvious change is, of course, the emergence of *ir 'go* as a variable context in the 20th-century spoken data; before this data set, there were
no examples of *ir a + ir 'going to + go' found in the data. Changes can be seen in other verb classes as well. While most verb classes show low percentages with both the PF and the SF, and are as such not as impressive, the pattern shown with stative verbs is unmistakable. While these verbs have made up a consistent (approximately) 30% of SF usage over the past seven centuries, they have had only minimal presence with the PF. Instead, we see a slow rise: they do not occur with PF until the 17th century, at a mere 5%. This climbs to 8% in the 19th and 20th-century written data sets, and nearly catches up to the SF rate (30%) in 20th-century spoken data, at 20%. Thus, while it appears that the distribution of verb classes remains steady over time for SF, PF gradually spreads to more classes.

Other early constraints on PF-SF variation, however, do not show the same pattern. If we return to Table 5.47, we see that the presence of a temporal adverbial, a significant factor in the 17th century, does not achieve statistical significance in the 19th century, arguably a victim of the generalization of the PF. This factor returns, however, in the 20th century. If we consider spoken language to be somewhat representative of more innovative language than written language, which is generally more conservative (Biber 1995), then we observe not a classic weakening of the adverbial constraint, but rather a strengthening. The same phenomenon occurs with sentence type, which exerts significant influence on PF-SF variation only in the 20th century. In line with Poplack and Malvar's interpretation of very similar results in Brazilian Portuguese (forthcoming), I believe that these newly significant factor groups have to do with the establishment of the PF as the default future, a status that is demonstrated in patterns occurring outside the envelope of variation discussed in this chapter.
5.3.5.1 Subject: measuring volition and intention

I would like to take a moment to discuss a factor group that was excluded from the VRA due to interactions with other factor groups: subject. There are two (related) semantic questions that an examination of grammatical person and subject animacy may help to clarify. First, as mentioned in Section 1.1.2, both the PF (e.g. Bishop 1973:89) and the SF (e.g. Fernández Ramírez 1986:284-285) have been associated with volition meaning. One way to test the potential for volition meaning is through subject animacy. Animacy is useful here because, in order for a subject to have volition, it must be animate (Bybee & Pagliuca 1987:114); the higher the rate of animate subjects occurring in a construction, the more speakers may be using it to express volition.

In general, as shown in Table 5.49, the PF has tended to occur more often than the SF with singular, animate subjects. In this case, the difference between the PF and the SF grew after the 15th century, after which there was a decrease in singular animate subjects with SF. The PF, on the other hand, with an oscillating rate of singular animate subjects, between 50% (6/12) and 68% (51/75), shows no significant differences or systematic patterning. There is a decrease in the rate of co-occurrence of singular animate subjects with the SF, a form that has been increasingly associated with inanimate subjects since the 15th century. Over the past 500 years, the rate of singular animate subjects with

---

104 The factor group of subject interacted with verb type, since inanimate subjects rarely occurred with psychological and perception verbs, and instead co-occurred with stative verbs.

105 This discussion is limited to singular subjects in order to match more closely the notions of volition and intention, which we would expect to be expressed more for singular subjects than for plural due to the internal and psychological nature of these notions.
the SF has dropped significantly, from 71% to 44% (*p* < .000, *x*²=82.69893), which suggests that the SF may be losing its ability to express volition.

Table 5.49. *Rate of singular animate subjects in PF and temporal SF by data set*

<table>
<thead>
<tr>
<th></th>
<th>PF</th>
<th>%</th>
<th>SF</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>1/1</td>
<td>100</td>
<td>621/1058</td>
<td>59</td>
</tr>
<tr>
<td>15</td>
<td>6/12</td>
<td>50</td>
<td>547/770</td>
<td>71</td>
</tr>
<tr>
<td>17</td>
<td>37/59</td>
<td>63</td>
<td>777/1270</td>
<td>61</td>
</tr>
<tr>
<td>19</td>
<td>51/75</td>
<td>68</td>
<td>249/432</td>
<td>58</td>
</tr>
<tr>
<td>20w</td>
<td>53/79</td>
<td>67</td>
<td>113/212</td>
<td>53</td>
</tr>
<tr>
<td>20s</td>
<td>453/765</td>
<td>59</td>
<td>171/392</td>
<td>44</td>
</tr>
</tbody>
</table>

This subject animacy pattern may be, in fact, a reflection of loss of intention meaning, expressed most easily through first-person uses. At this point, whether "intention" or "volition" is the most adequate characterization for earlier SF meanings is unclear; volition (as operationalized here) does not appear to hold much explanatory power, however, in the case of the PF. However, as we will now see, an examination of first-person uses in operationalization of intention meaning shows relatively robust results for both constructions.

In Table 5.50, we see that the use of first-person singular subjects with temporal SF and PF show opposite diachronic tendencies. First, with the SF, there is a decrease in first person singular between the 17th and 20th-century written data, from 36% (453/1270) to 20% (42/212), suggesting a loss of ability to express intention during this time period. At the same time that the SF first-person uses decline, the PF begins to be used more with the first person singular, from 27% (16/59) to 40% (30/75), suggesting that the PF's ability to express intention increased. What we see here, then, is a reversal in tendencies.

---

106 In New Mexican Spanish, Villa-Crésap (1997:56) reports a low rate of 7% intention uses for the SF, compared to 27% for the PF (1997:64).
over time as speakers come to associate the PF with intention meaning. Though overall, totaling results from all centuries, neither form shows a general tendency to occur more in first person singular than the other (the average proportion of first person singular in temporal SF for the data as a whole is 30% (1241/4121), and for the PF it is 31% (306/994)), the 1s SF percentage is higher until the 17th century, after which the 1s PF percentage is higher.

Table 5.50. Rate of first-person singular subject in temporal SF and PF by data set

<table>
<thead>
<tr>
<th></th>
<th>OldSp % (N)</th>
<th>17 % (N)</th>
<th>19 % (N)</th>
<th>20w % (N)</th>
<th>20s % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF 1s</td>
<td>30 (541)</td>
<td>36 (453)</td>
<td>29 (124)</td>
<td>20 (42)</td>
<td>21 (81)</td>
</tr>
<tr>
<td>PF 1s</td>
<td>23 (3)</td>
<td>27 (16)</td>
<td>40 (30)</td>
<td>35 (28)</td>
<td>30 (229)</td>
</tr>
</tbody>
</table>

SF: 17th vs. 19th, p ≤ .0082, x² = 6.979659; 19th vs. 20w, p ≤ .0153, x² = 5.877678
PF: 17th vs. 19th, p ≤ .0801, x² = 3.0625
No other significant differences between data sets.

These simultaneous changes in the rate of first-person singular subjects do not explain, however, why attempts at operationalizing the notion of intention through coding for first person singular have given unsatisfactory results in variationist studies (e.g. Sedano 1994; Poplack & Turpin 1999; Poplack & Malvar forthcoming), since it appears (as suggested in the descriptive literature) that both forms have been associated with intention meanings, though at different times. For instance, pair-wise comparisons show that the difference between PF and SF in 20th-century speech reaches statistical significance in this corpus, p = .0024, Chi-square = 9.177948, as does the difference in 20th-century plays, p = .0055, Chi-square = 7.69802. The association of both constructions with intention at different times is expected if we take into account Bybee, Pagliuca and Perkins' (1991) models for future form development, which posit intention meaning for modal-based forms, like the SF, and possibly for motion-based forms, like

207
the PF. Unfortunately, the diachronic evidence does not seem to answer the ever-present question of which came first, the innovative PF form's pragmatic strengthening (Traugott & Hopper 1993) or the conservative SF form's semantic weakening. Both processes seem to occur simultaneously, beginning around the 17th century.

We may attempt to operationalize the notion of prediction in the use of these constructions with inanimate subjects, which do not, under normal (non-figurative) conditions, have intentions. According to Bybee, Pagliuca and Perkins' (1991) models for future development, prediction meaning would follow from intention; we would expect to see a rise in inanimate subject rates as this part of the grammaticization process took place. Table 5.51 shows the proportion of PF and SF uses that have occurred with inanimate subjects by data set. Here, the overall rates reveal that inanimate subjects have accounted for a significantly higher proportion ($p < .0000$, Chi-square = 17.29395) of the SF at 17% (703/4121) than of the PF at 12% (116/994), and the SF inanimate rate has been higher than the PF rate in every data set (Old Spanish has too few PF tokens for comparison), suggesting that, in general, the SF has been used more to express prediction than the PF, a finding that may explain in part the perception in some of the linguistics literature that the SF is a more "neutral" or "objective" future (cf. summary of literature in Poplack & Turpin 1999:137).

A look at the diachronic data shows that the PF shows no significant differences in singular inanimate subject rates between data sets except for in the 17th century, which showed no singular inanimate subjects. The highest rate for inanimate subjects with the PF occurs in the 19th century data, at 17% (13/75), which could indicate an increase in
prediction meaning. After this, the rate drops, though not significantly, to 10% (8/79) in 20th-century writing and 12% (93/768) with the PF.

Table 5.51. Rate of singular inanimate subjects in temporal SF and PF by data set

<table>
<thead>
<tr>
<th>Subject</th>
<th>OldSp</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF 3s inanimate</td>
<td>14 (265)</td>
<td>18 (230)</td>
<td>18 (77)</td>
<td>17 (37)</td>
<td>25 (94)</td>
<td>17 (703)</td>
</tr>
<tr>
<td>PF 3s inanimate</td>
<td>15 (2)</td>
<td>--</td>
<td>17 (13)</td>
<td>10 (8)</td>
<td>12 (93)</td>
<td>12 (116)</td>
</tr>
</tbody>
</table>

SF Total N = 1828 1270 432 212 379 4121 13 59 75 79 768 994

Difference OldSp-17 for SF significant ($p = .0069, \chi^2 = 7.288573$), and 20w-20s for SF significant ($p < .0000, \chi^2 = 65.5746$); no other differences significant

With the SF, we see a significant increase in co-occurrence with inanimate subjects between Old Spanish and the 17th-century Peninsular Spanish data ($p = .0069, \chi^2 = 7.288573$) and again in 20th-century speech ($p < .0000, \chi^2 = 65.5746$). The first change seen here may be a reflection of SF's early extension into more prediction contexts and subsequent loss of intention meaning in the 17th-19th centuries (Table 5.50). The 20th-century difference may be due to a rise in SF use with prediction meaning both within and outside future temporal expression (i.e. as an epistemic marker) during the 20th century, since epistemic uses of SF are most likely to co-occur with singular inanimate subjects (Section 3.4.2). Indeed, the epistemic uses may have developed with the SF because speakers already used the SF to express prediction more often than they used the PF with this meaning, an association that became stronger over time, perhaps—ironically, given the centrality of prediction to future meaning—eventually causing the SF to lose default status within the realm of future expression.
In this section, I have attempted to operationalize the notions of volition, intention, and prediction in terms of subject animacy and person.\textsuperscript{107} It was shown that the rate of singular animate subjects, meant to operationalize volition, did not show systematic results (Table 5.49), suggesting that volition may not be relevant to SF-PF variation. The operationization of intention through first-person singular, however, revealed that, beginning in the 17\textsuperscript{th} century in these Peninsular data, the SF underwent a significant decrease in first-person singular subject co-occurrence, while the PF experienced a first-person singular rate increase during the same period (Table 5.50), which I took as evidence for loss of intention meaning in SF and a rise of the use of PF in contexts with intention readings. Finally, an examination of the rates of singular inanimate subjects (Table 5.51) as a measure of prediction meaning revealed significant increases in SF, though no significant changes were found in PF.

5.4 Summary and discussion

5.4.1 Semantics and variant choice: a diachronic perspective

In this chapter I have used variationist methodology to uncover the linguistic conditioning of the choice between Spanish PF and SF as futures since the 17\textsuperscript{th} century. Many of the factors examined were expected to shed light on the possible semantic motivations conditioning the choice between PF and SF. The semantic notions operationalized in this study include:

i) retention of motion meaning, through verb class

\textsuperscript{107} This factor group was considered separately from the variable rule analyses due to statistical interactions with verb class.
ii) temporal modification and temporal specificity, through temporal adverbials

iii) (un)certainty, through sentence type and head verbs of subordinate clauses

iv) volition, through subject animacy, and

v) intention, through grammatical person.

Other factors examined were clause type, based on Bybee’s (2003) assertion that main clauses are more innovative; and polarity, due to Poplack and Turpin’s (1999) finding that this was the strongest factor influencing PF-SF variation in present-day Canadian French.

With regards to the five semantic notions explored, three showed a clear and consistent association with one construction; i) dynamicity, ii) lack of explicit temporality and, more weakly, a specific time frame, and iii) volition are and always have been associated more with the PF. It should be noted that these meanings are harmonic with the PF's lexical origins as "agent on a path toward a goal" (Bybee & Pagliuca 1987).

Conversely, stativity and nonspecific time in particular have been consistently associated with the SF, which are, in fact, consistent with the lexical aspect of its lexical source Infinitive + habere 'have', despite the fact that any original possessive meaning this construction may have had was lost before the birth of Spanish (see Section 3.1). The SF showed a loss of volition meaning over time, as the rate of animate subjects declined.

The other two meanings, certainty and intention, showed more complex results. I attempted to operationalize the notion of "certainty" in two specific sites: sentence type and the semantics of heads of object complement clauses in the 19th century and in 20th-century speech. In complement clauses, a clear association of the PF with certainty was found, as it occurred most often in subordinate clauses headed by verbs of certainty such
as saber 'know' or belief, like creer 'believe', unlike the SF, which occurred most with verbs of uncertainty like no saber 'not know'. This tendency was seen in both the 19th century and the 20th-century spoken data sets. Another trend that was revealed here was a 19th-century association of the SF with contexts of "promise," an association that was absent from the 20th-century spoken Peninsular Spanish.

The results for sentence type did not corroborate the findings in complement clauses; there has been a consistent favoring of the PF in interrogative contexts, which were hypothesized to correlate with uncertainty meanings. These results suggest, I believe, that perhaps interrogative contexts are not as good of an operationalization of uncertainty meaning as I had hoped. It is doubtful, given the robust and more transparent nature of the complement clause findings and the conclusions of previous scholars, most of whom have associated the PF with certainty (e.g. Bishop 1973:89; Confais 1995; Villa-Crésap 1997; Almeida & Díaz 1998; Jensen 2002), that the PF should be associated generally with uncertainty.

The semantic notion of intention is the feature that showed the most complex transformation, which is probably why synchronic quantitative studies (e.g. Sedano 1994; Poplack & Turpin 1999) have been unsuccessful in associating rates of first-person singular subjects with intention. The diachronic data show that there was a simultaneous increase in first-person use with PF in the 17th century and a decrease in

108 Sedano, associating the PF with intention, does so not through quantitative analysis of first-person singular rates, but rather through intuitive interpretation of first-person uses: "La única alternativa para medir la intención del sujeto en la emisión del FMT [SF] o del FP [PF] es guiarnos por nuestra intuición de hablantes nativos del español y observar atentamente los casos del corpus en los que esas formas aparecen conjugadas en primera persona del singular" (1994:235).
first-person subject use with the SF, suggesting that the PF moved into intention meaning as the SF was losing its ability to express intention.

The two syntactic factors that were not associated with semantic notions, clause type and polarity, did not line up with the findings of Bybee (2003) regarding clause type or Poplack and Turpin (1999) regarding polarity. Bybee (2003) proposes that main clauses are innovative and subordinate clauses are conservative; I thus hypothesized that the innovative form, the PF, would be favored in main clauses. This was not the case; the PF has consistently been favored in subordinate clauses, significantly in the 19th century. Poplack and Tagliamonte (2001) found the same direction of effect, with subordinate clauses favoring going to, in all varieties of English they examined, an effect that reached statistical significance in two of the five varieties in their study (2001:227-228).109 Negative polarity, which strongly favors SF in present-day spoken Canadian French (Poplack & Turpin 1999), was also found to favor the SF in Peninsular Spanish through 20th c. written data (significantly in the 17th century); in 20th-century spoken Peninsular Spanish, this direction is reversed, and negative polarity significantly disfavors the SF.

---

109 Poplack and Malvar attribute this effect to a persistence of the volition meaning of will: "Going to, with purportedly less volitional import, is said to occur more frequently in subordinate clauses (Royster & Steadman 1923/1968:400). As grammaticization proceeds, the contribution of clause type, as an instantiation of both the original point of entry of going to, and the persistence of the volitional meaning associated with will, should decrease as we proceed from the more conservative enclaves to the mainstream variety" (2001:228). Such an explanation would not apply here, however, since the SF did not originate from a construction expressing volition.
5.4.2 (In)stability and renewal

One of the goals of this chapter was to examine empirically the local process of generalization, i.e., what path, and in what contexts, did the PF take during semantic and contextual generalization? In the face of radical frequency change, one notable feature of the past 300 years of PF-SF variation is that the direction of effect in all factor groups examined except polarity has remained constant. This shows that, even nearly half a millennium after a construction begins to grammaticize, quantitative analyses can still show evidence of a construction's original (lexical) meaning. Changes in relative frequency, then, do not necessarily imply similarly drastic changes in the internal structure of variation. Likewise, we may consider that, during the grammaticization of a form, the subtle changes that emerge in a variety's grammar may take place at a slower pace than the more noticeable changes in frequency (cf. Poplack & Malvar forthcoming).

At the same time, nonetheless, in the case of the Spanish PF, it appears that not all disfavoring contexts are equal. In the 17th century, we saw that the PF was disfavored with stative/psychological/perception verbs, when there was a temporal adverbial, and in negative polarity contexts. Each of these factors showed a different magnitude of effect on variant choice; verb class was the strongest. This factor, over 300 years later, is the only one that remained consistently significant in all data sets, losing magnitude of effect as the PF generalized into new functions. The factor of temporal specificity lost significance in the 19th century, and regained significance in the 20th century; I will propose, however, that this return to significance was neither an accident of the statistical analyses nor brought about by strong disfavoring of PF because of a return of PF to lexical semantics, but rather by reinforcement of the favoring of SF with nonspecific
temporal adverbials because of the development of new non-future uses of SF (see the discussion in Section 6.4.3). The factor of polarity was, in fact, reversed: the PF is favored in negative polarity in 20th-century Peninsular Spanish speech.

This portrait of shifts in conditioning on PF-SF variation is a reflection of two simultaneous processes: first, the semantic and contextual generalization of the PF, and second, the semantic generalization of the SF in the 20th century. In Section 5.1.2, I made three hypotheses: i) the PF would be favored with animate subjects; ii) the PF would be favored with specific temporal adverbials; and iii) the PF would be disfavored with stative and psychological verbs, and with ir 'go'. Regarding the first hypothesis, it was shown that, in general terms, the PF has had a higher rate of animate subjects than the SF, though this difference fails to reach significance except when the entire corpus is taken as a whole. More revealing is the fact that, while there has been no significant change in the rates of animate subjects with the PF, the SF has shown a significant decrease in the relative frequency of co-occurring animate subjects (Section 5.3.5.1, esp. Table 5.51).

These results, then, cannot be taken as evidence of PF generalization, since it was in fact the contexts in which the SF occurred, and not PF contexts, that changed.

Regarding the second hypothesis, that PF would be favored with specific temporal adverbials (when an adverbial was present), is borne out in the three data sets in which specificity could be measured (the 19th century, and the 20th-century written and oral), though the context in which speakers have most preferred the PF has consistently been without any temporal adverbial modification whatsoever (see Table 5.45 and Table 5.46). I will return to some possible interpretations of these patterns in Section 6.3.3.
The third hypothesis, that the PF would be disfavored with verb classes inconsistent with purposive motion meaning, was also shown to be correct. The PF has been consistently disfavored with stative, psychological, and perception verbs in all data sets, and categorically absent with the verb *ir* 'go', until the 20th-century spoken Peninsular Spanish data. This factor group, with the most robust effect in the early stages of development, has gradually diminished in effect since the 17th century, though the direction of effect has remained constant (see Table 5.45 and Table 5.46), presenting a nice case of contextual generalization.

The results presented here reveal that PF generalization did not happen at the same pace across all contexts. The PF was historically disfavored in three contexts: stativity; temporal adverbial modification and, where modified, with nonspecific temporal adverbials; and negative polarity. These effects dissipated at different rates across different contexts: in negative polarity first, then in adverbial modification (at least for a short time, before the SF began its new process), and last, still holding on even in 20th-century Peninsular Spanish speech, stativity. This diachronically uneven incorporation of the PF into erstwhile disfavoring contexts throws doubt on the validity of Kroch’s (1989) Constant Rate Hypothesis, since the PF did not emerge in all contexts at the same rate (see Poplack & Malvar forthcoming for the same conclusion based on evidence from Brazilian Portuguese).

The evidence of this variable rate, seen in changes in relative magnitude of effect and loss of significance as constraints weaken, is further complicated by something else that appears to be happening in the 20th century. Other changes we see, namely a strengthening of the temporal specificity and sentence type constraints, I believe to be
related to the further grammaticization of the SF in the 20\textsuperscript{th} century as it emerged into epistemic uses, the fourth and final stage of grammaticization for modality-based future forms proposed by Bybee, Pagliuca an Perkins (1991). This second process, which has ramifications that extend beyond the variable context studied in this chapter, will be the focus of Chapter 6.
6 Synchronized change

6.1 Unresolved questions

6.1.1 Beyond the variable context

So far this study has relied heavily upon both grammaticization theory and variationist methodology, following most closely recent work of Shana Poplack and colleagues on future expression that have also drawn on both approaches (e.g. Poplack & Turpin 1999; Poplack & Tagliamonte 1999, 2001; Walker & Torres Cacoulos 2005; Poplack & Malvar forthcoming). Nevertheless, my analysis of the diachrony of Spanish futurity presented in Chapter 5 left various questions unanswered. For example, the 20th century shows increased magnitude of the adverbial effect, with co-occurring temporal adverbials, especially non-specific ones, favoring SF over PF. Another unexplained change seen in the 20th century is a growing favoring of PF in interrogative contexts.

A search for answers has led me down an unexpected path, which combines the form-based study so essential to the study of grammaticization and the function-based analyses which are the key component of variationist studies. In Chapter 5, we explored and confirmed the hypothesis that the lexical origins of the newer PF would be reflected in the diachronic trajectory of constraints on its variation with the SF, the latter of which is shown to be highly grammaticized (cf. Bybee, Pagliuca & Perkins 1991, whose model would place the 20th-century SF in the last of four stages in the grammaticization of modal-derived futures, due to its epistemic uses). It is simultaneously recognized that fierce competition can lead to restriction, fossilization, or even death of conservative forms (Bybee, Perkins & Pagliuca 1994; Company 2003:44). In Portuguese, for example,
the SF has simply disappeared from usage as a future even in educated speech (Poplack & Malvar forthcoming); its use (if any is to be found) outside of future contexts has yet to be analyzed quantitatively. In Peninsular Spanish, the future SF has not reached such a dire state (though studies suggest that the restriction to non-future epistemic uses is more advanced in American Spanish; Boyd-Bowman 1960:166; Moreno de Alba 1978:89; Cartagena 1995-1996; Villa-Crésap 1997:58). Given the future SF's demise in sister varieties, it seems quite reasonable to fancy the notion that a similar fate awaits the SF of Peninsular Spanish.

While a straightforward application of grammaticization theory and the variationist framework allowed me predict most of the patterns we saw in Chapter 5, such as increased PF (relative and raw) frequency and weakening of constraints tied to lexical origins, other patterns have been left unexplained. Like Poplack and Malvar (forthcoming) found in Portuguese, unpredicted changes in the conditioning on variation were observed in the 20th-century data sets. For instance, if the PF is generalizing, and has in fact reached the status of generalized future in the 20th century (Melis 2006), and generalizing forms "inicia[n] en contextos muy marcados o específicos y progresivamente pierde[n] restricciones de empleo y se generaliza[n]" (Company 2003:26), then we might expect to see a progressive weakening over time of constraints on PF-SF variation that reflect the lexical origins of the most innovative form in question.

While we do see such patterns in the 20th century, namely in verb class constraints, which weaken over time as speakers gradually use the PF more with stative, psychological and perception verbs, we also see trends indicating strengthened constraints. These trends include: i) a strengthening of the temporal specificity constraint,
shown in an increase in its relative magnitude of effect; and ii) an increase in the magnitude of effect of the factor of sentence type, leading to its selection as statistically significant. Even if the constraints shown to be significant in the 20th century could be plausibly explained in terms of the lexical origins of the SF or PF, the fact that they have become stronger over time throws such an argument into doubt.

While a focus on the space of competition, future temporal reference, as in the previous chapter, is quite valuable in describing the nature and direction of a change, and in understanding some of the conditioning of structured variability, it is not sufficient in this case to explain all the shifts in the conditioning of PF-SF variation that emerge in the 20th century. Instead, we must apply the results of variationist methodology to the notion of grammaticization in a rather unconventional way.

I will argue here that the territory into which the PF was generalizing became perturbed in the 20th century: as the PF began to take over SF territory, the SF itself continued to grammaticize as well. Whether this upset was internal, i.e. evoked by the intrusion of the PF, or external, i.e. evoked by the SF as it lost semantic content as it went along its merry way in grammaticization, a closer look at the results of form- and function-based studies uncovers some suspiciously similar phenomena inside and outside the envelope of variation of future temporal reference.

Almeida and Diaz (1998), Poplack and Turpin (1999), Poplack and Tagliamonte (2001) and Poplack and Malvar (forthcoming), in their studies of Spanish, French, English and Portuguese futures, respectively, do not consider the possibility that non-future epistemic uses of the French or Portuguese Synthetic (or Morphological) Future or English will might affect patterns of variation in temporal future expression. Epistemic
uses of future forms are well documented in all four languages (on English, Comrie 1985:43-48; Arnovick 1990; Vettes and Skibinska 1998; on French, J. Lyons 1968:310; Tomaszkiewicz 1988; on Portuguese, Tlaskal 1978:212-213), though natural examples are somewhat difficult to find in the literature. A constructed French example, taken from Palmer (2001:105), is seen in (6.167), and an example from Fleischman (1982:132, trans. by Fleischman), is seen in (6.168).

(6.167) Ça sera le facteur That be+3SG+FUT+IND the postman 'That'll be the postman' (epistemic)

(6.168) George n'est pas venu ce matin. Il aura oublié notre rendez-vous. 'George didn't come this morning. He probably forgot [lit. will have forgotten (SF)] our appointment.'

Of course, it is likely a search outside the future temporal reference sector would have made little sense, at least in the case of Canadian French, in which Poplack & Turpin (1999:161) note that non-future uses of SF are "exceedingly rare" in their spoken data; Poplack and Turpin do mention that non-future uses of SF in Canadian French amount to about 20% of SF data in their spoken corpus (1999:160). However, unfortunately, little is known about the rate or productivity of such uses in the Brazilian Portuguese variety studied by Poplack and Malvar (forthcoming) and of the Canary Islands Spanish variety

---

110 Poplack and Turpin (1999:160) offer the following as justification for the lack of attention paid to SF outside future temporal reference contexts: "Indeed, despite its rarity in future temporal reference contexts, IF [i.e. SF] is actually employed four times more often for this purpose than for any other. This should allay any suspicion that IF fulfills a major grammatical role outside the future temporal reference sector." While the rate of about 20% may seem high, since it parallels the rate of 24% in 20th-century spoken Peninsular Spanish data in the present study, the overall frequency of SF in their study was much lower than the frequency found in the data presented here.
studied by Almeida and Díaz (1998). Furthermore, a handful of non-variationist studies of Spanish offer quantitative information regarding this use. Sedano (1994), for example, found epistemic uses of SF made up 58% (148/249) of SF use in her Venezuelan Spanish data, and Moreno de Alba (1978:98) found 73% uncertainty use. Villa-Crésap (1997:58) found a rate of 70% "uncertainty" in New Mexican Spanish SF, and Gradoville and Durán Urrea (2006) reported 79% "modal" uses of SF. Thus, the vitality of epistemic SF may be much greater in Spanish (and Portuguese) than in French.

Poplack and Turpin, in their study of French, cite low relative frequency as reason to follow what they characterize as "the standard variationist practice of excluding tokens which do not form part of the variable context," while simultaneously acknowledging the possible objection that this practice "effectively obscures any productive role IF [i.e. SF] may play elsewhere in the grammar" (1999:160). My results will show that, in the case of Peninsular Spanish future, to disregard epistemic uses of the SF—nearly a quarter (24%) of 20th-century SF use in speech—without at least minimal examination, would be to discard a tremendous amount of explanatory power. I will show here that the epistemic SF and conditioning of SF-PF variation in future expression appear to have a secret relationship.

In this chapter, I will attempt to take full advantage of variationist methodology and grammaticization theory, which offer the "posibilidad de integrar o vincular fenómenos y datos que tradicionalmente [han] estado desvinculados" (Company

---

111 The rate reported by Villa-Crésap (1997:58) measures what he identified to be "low-certainty contexts," which included future temporal reference; thus this figure does not measure the exact same phenomenon I am referring to here.
by presenting various pieces of evidence that reveal an intimate connection between the patterns of variation found in 19th- and particularly 20th-century Spanish future temporal expression, and the path of grammaticization followed by the Synthetic Future. A comparison of constraints on PF-SF variation in future expression and distributional patterns of epistemic SF shows parallel changes in overall rates of occurrence, sentence type and verb class effects, as well as in the rate of temporal adverbials. Some of the changes seen in the 20th century suggest that the temporal SF

Another change we see is that the PF becomes the most common form in negative polarity contexts (see Section 5.3); epistemic SF uses tend to occur with positive polarity, as shown in Table 6a. Here we see that 6% of positive polarity contexts in SF in all of the data are epistemic, and only 3% of the negative polarity uses of SF are epistemic. This is a small but significant difference evidenced in the data as a whole (p < .01). The polarity effect is also very weak in the conditioning of PF-SF variation in future temporal expression, with a small range of 12 in variable rule analysis results for the 20th-century speech data.

Table 6a. Distribution of temporal and epistemic SF by polarity, all data sets

<table>
<thead>
<tr>
<th></th>
<th>Temporal</th>
<th>Epistemic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Negative</td>
<td>599</td>
<td>97</td>
<td>17</td>
</tr>
<tr>
<td>Positive</td>
<td>3534</td>
<td>94</td>
<td>212</td>
</tr>
</tbody>
</table>

p ≤ .0028, $\chi^2 = 8.941664$

When it comes to the question of polarity, however, this type of parallel shift in the same direction is not evident. Here, the temporal SF is more like the epistemic SF, and less like the PF. A look at the proportion of negative polarity in each use, depicted in Table 6b, gives a clearer picture of this alignment.

Table 6b. Proportion of negative polarity, by data set

<table>
<thead>
<tr>
<th></th>
<th>PF</th>
<th>Temp. SF</th>
<th>Epistemic SF</th>
</tr>
</thead>
<tbody>
<tr>
<td>OldSp</td>
<td>0</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>7</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>20w</td>
<td>9</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>20s</td>
<td>11</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

As seen in Table 6b, it is only in 20th-century spoken data that we see a likeness of the temporal SF with the epistemic SF; in previous data sets, the proportion of negative polarity with temporal SF was double or triple the proportion of that found in PF and epistemic SF, both of which co-occurred with positive polarity.
and epistemic SF have actively diverged (see 1.1.2.2 on Comrie's proposal to distinguish epistemic and temporal uses of same form as two separate grams). Evidence for these claims is drawn from careful examination of the results of the form-based study of the SF presented in Chapter 3 and the complex results of the function-based analysis offered in Chapter 5.

6.1.2 Semantic generalization and functional space

In the case of the SF, the 20th century saw a rise in non-temporal epistemic uses. We may hypothesize, then, that some of the changes observed in the 20th century are not to be taken as evidence of contextual generalization of PF, nor of new functions within future expression, nor of retention of lexical meanings of either construction, but rather as consequences of the SF's loss of productivity in the realm of future temporal expression.

It is interesting to note that the PF use in negative polarity has slowly risen, showing a loss of the polarity constraint hypothesized to originate in the lexical origins of the PF. In 20th-century speech, the proportion of PF used in negative contexts reaches 11%, which is at the cusp of the typical rate in temporal SF of 12-18% (until the 20th-century spoken data). Like verb class constraints on PF occurrence, the polarity constraint seems to be almost lost in the final data set. We may consider this, then, a nice case of generalization of function.

Such a process of generalization, however, does not explain the reversal in the direction of effect seen in 20th-century spoken data, since, like stative verb use in PF, rates have still not reached the general SF rate of 12%-18%. As Table 6b shows, not only has the PF negative polarity rate risen to nearly meet temporal SF averages, but the temporal SF proportion in negative polarity drops from 18% in 20th-century written data to only 8% in the 20th-century spoken data. The epistemic SF, in turn, shows a slow but steady rise in negative polarity, in many ways mirroring the PF. Unlike in the case of interrogatives, where the temporal SF diverges from the epistemic SF in the 20th century, in the case of polarity the temporal SF—after hundreds of years of stability—follows suit with epistemic SF tendencies.
and its loss of status as default future (see Poplack & Malvar forthcoming for a similar argument for future temporal expression in Brazilian Portuguese). It is in the presence of incipient or moribund forms, Poplack and Malvar (forthcoming) note, that the internal stability of the division of constraints conditioning variation within a functional space may be "perturbed." In this case, we will see why this might be through an examination of SF use outside the variable context.

The loss of default future status for SF was not instant, nor has its association with futurity been lost completely; the SF still accounts for a full third (33%) of future expression even in 20th-century Peninsular Spanish speech, and contexts that favor an epistemic reading of SF only account for one-fourth (24%) of SF usage. Let us put aside once again uses that fall outside the variable context of futurity. If there are shifts in the division of labor between PF and SF, and these shifts cannot be adequately explained in terms of the generalization of the PF or the SF within future expression, then we may hypothesize that the division of labor between variants has been somehow perturbed. As Poplack and Malvar show, this can happen when a newer for gains default status (forthcoming). In the case of Spanish future temporal expression, in which we know that one variant (the SF) is performing functions outside the variable context, a search of the behavior of these forms outside the variable context may yield clues about the driving force behind these changes. Since the use of PF in contexts that do not allow future

113 Note that this is not the same situation described in Bybee, Perkins & Pagliuca (1994:230-236, 296), in which erstwhile indicative forms with no modal meanings take on modal meaning through the absorption of the linguistic context in subordinate clauses. The SF has always had modal meaning since it began to be used as a future, and epistemic SF—the innovative use of interest here—tends to occur in main clauses (see Section 3.4.2).
readings is exceedingly rare in these data, I have turned to the use of SF in epistemic contexts, the most frequent use of these forms without possible future readings, for possible clues.

As we saw in Chapter 3, however, the search for notable change in SF, especially when it is compared to the striking patterns of generalization and rise in frequency of the PF (see Chapter 4), seems less than promising: an examination of the SF distribution patterns reveals a (mostly) stable temporal SF, which has occurred in about the same contexts at about the same rate for hundreds of years (see Section 3.4.1). At no point do we observe in the temporal SF the same magnitude of change in frequency and distribution that we observe in the temporal PF. There are, nonetheless, some small but significant changes in temporal SF distribution in the 20th century. In particular, I would like to address here the decrease in the rate of interrogatives in temporal SF (Section 3.4.1) and a rise in the rate of nonspecific temporal adverbials (see Section 6.4.3 below).114

In Portuguese future expression, Poplack and Malvar's (forthcoming) results showed an intricately woven functional space that remained internally stable; however, they also found that constraints were transferred from form to form as new forms moved in and older forms moved out of the functional space, for example, the contingency constraint, associated with the SF, was transferred to P when the SF effectively disappeared. The transferal of constraints from one form to another results in

114 There were also significant changes in SF patterns according to subject person and animacy (see Section 5.3.5.1). However, since this factor was not included in the function-based statistical analyses in Chapter 5, and thus the relative magnitude of effect of this factor is unknown, grammatical person and subject animacy cannot be adequately addressed here.
strengthening or even reversal of constraints on the choice of a particular form over time, similar to the unexplained changes seen in 20th-century Peninsular Spanish future expression.

Poplack and Malvar's (forthcoming) study, which included four future forms, portrays a complex space in which the division of labor is not simply a smooth transition from old to new form, but rather a complex balancing act with four main actors. Since my study is limited to only two future forms, I unfortunately cannot provide the same level of complexity in my analysis regarding the stability of the realm of Spanish future expression in face of other competing constructions, particularly the futurate Present. Nevertheless, their findings inside the realm of Brazilian Portuguese futurity, particularly that incipient and moribund forms can perturb the division of labor between variants for particular forms in variation, appear to apply to the Peninsular Spanish future as well. In this case, however, the key to understanding the particular contexts for change in the 20th-century lies outside the realm of future temporal expression. It appears here that the face of 20th-century Spanish future is altered by the shift of the SF from more tense-like to more mood-like. If we assume a temporal SF that is losing productivity as predicted in grammaticization theory (e.g. Company 2003:50), then new questions are raised: If epistemic uses represent the loss of default future status for the SF, why is the SF losing or gaining ground in particular temporal contexts and not in others? Which shifts in strengths of constraints on variation (temporal adverbials, sentence type, verb class) can be attributed to PF contextual generalization, and which to a growing tendency to use SF in contexts without future temporal reference?
6.2 Frequency

I have suggested the possibility that the path of grammaticization and generalization of the PF within the realm of future temporal expression, the shifting strength of constraints on PF-SF variation, and the semantic change in the SF outside this realm are somehow linked. The first clue that points to such a hypothesis is a simple diachronic observation: the epistemic use of the SF surged in relative frequency at about the same time that the PF surged in relative frequency within future temporal expression.\textsuperscript{115}

Figure 5.6. Relative frequencies of temporal PF and epistemic SF by data set

Figure 5.6 shows the relative frequencies of the epistemic SF (relative to temporal SF) and temporal PF (also relative to temporal SF). As we have seen in Chapter 3, the

\textsuperscript{115} It would be interesting and informative to examine the status of other constructions expressing epistemic modality in Peninsular Spanish during this time, in order to more fully understand the space into which epistemic SF was emerging. Unfortunately, variation within the realm of epistemic modality is beyond the scope of the present study.
epistemic use of the SF was nearly imperceptible throughout the entire period examined—at only 2%—until the 20th century. As this figure shows, the rise in epistemic use of the SF parallels the rise in use of the PF as a future, with a lag of about a century. While the PF makes up about a tenth of the data (12%) in the 19th-century texts, epistemic SF reaches the same prominence (10%) in the 20th-century texts. Similarly, the PF makes up a quarter of the data (27%) in 20th-century texts, while it is in 20th-century speech that the epistemic SF makes up 24%.

This figure, at the very least, shows that the rise of these two form-meaning pairings moved in parallel fashion. Such a pattern raises an intriguing question: is this evidence that semantic competition (between PF and SF as future expressions) can, in fact, trigger further grammaticization of an erstwhile stable form (SF)? The diachrony in Figure 5.6 suggests that this may be the case; that is, it seems that as the PF moved into viable competition with the SF, it began pushing the SF out of the realm of competition (futurity) and into another realm (epistemicity). In other words, it appears that the SF is getting knocked out of the ring as it loses its ability to express intention (see Section 3.4.1).116

The validity of the notion of a connection between the advancement of the grammaticization of these two forms during their intimate 300-year negotiation of Spanish futurity can be evaluated by turning to variationist methods. This methodology

---

116 This question begs for cross-linguistic data. If this is, indeed, an interconnected change, then the patterns that emerged here should differ from those of languages whose go-future competes with forms that differ from the SF in lexico-semantic origins (non-modal-based futures). If, however, this is simply a case of persistence, these patterns should approximate those of other go-futures, without regard to the nature of competing forms. I would like to thank Rena Torres Cacoullos for pointing out this prediction.
provides the fruits of quantitative analysis in such a way that, if we wish, we can go beyond the comparison of forms in competition within the same semantic space. In this case, we are interested in two forms that are intimately linked in a different sort of way. The nature of this relationship, which has implications for grammaticization theory and our approach to variation, is the focus of this chapter.

6.3 Changing magnitudes of effect

For the clearest portrait of futurity, we will return to the variable rule analyses in Chapter 5. During the 20th century, as the epistemic use of the SF was on the rise (see Section 3.4.2), multivariate analyses reveal changes in the internal landscape of Spanish future expression. These changes take two forms. First, we see a change in the ordering of the constraints conditioning PF-SF variation, through the weakening of the strongest constraint in the history of this variation, verb class, as the PF generalizes into stative contexts. Second, we see tendencies in distribution that were previously not selected as statistically significant, such as a favoring of the PF in interrogative contexts, reaching a higher magnitude of effect than in previous data sets. These strengthened factor groups included sentence type and the presence of a temporal adverbial (previously weakened in the 19th century).

As these shifts in distribution took place within future temporal expression, speakers began to use the epistemic SF with more appreciable frequency. Certain contexts show significant differences between temporal and epistemic SF use in the 20th-century speech data, namely, epistemic uses of the SF tend to occur in interrogative contexts with stative verbs and no temporal adverbial (Section 3.4.2). It is interesting to
note that these contexts are related to the same contextual features that underwent changes in 20th-century PF-SF variation: verb class, sentence type, and temporal adverbials (see Chapter 5).

In the first two, the changes move in what may be considered the same direction: the PF generalizes into stative contexts and the epistemic SF occurs in mostly stative contexts; and both the PF and the epistemic SF occur more often in interrogative contexts than the temporal SF. In these two factors, then, the forms' distributional patterns mirror each other, one inside the realm of future temporal expression, the other outside (in the outskirts, if you will). In the case of temporal adverbials, a less direct relationship is seen: the rise in epistemic SF use weakens the association of the SF with futurity, which is accompanied by a 20th-century rise in temporal adverbials co-occurring with temporal SF (thereby strengthening the PF's already established tendency to occur without a temporal adverbial). In this section, I will describe the PF and epistemic SF patterns found in each of these three contextual features in an attempt to uncover the systematicity in the distributions of 20th-century PF and epistemic SF, whose similarities seem too impressive to be coincidental.

6.3.1 Stative verbs

Verb class is the one constraint on 20th-century PF-SF variation that appears to be the last stronghold of the lexical origins of the PF. This factor group, which means the disfavoring of the PF in contexts inconsistent with motion, such as psychological and stative verbs, loses its status as strongest constraint in 20th-century speech. As we saw in Table 5.48, repeated below as Table 6.52, stative verbs have remained at around 30% in
the SF since the 13th century, while stative verbs slowly entered the PF beginning in the 17th century with only 5% of PF occurrences. Only in 20th-century speech does the PF come close to the SF’s average, at 20%. Nevertheless, stative verbs still favor SF, as 45% of statives occur with SF, above SF’s average frequency of 33% in 20th-century speech in general (Table 5.42); this favoring is shown in the probability weight of .36 for PF in stative contexts (Section 5.3.4), showing a disfavoring of PF. This is an example of a classic case of (as of yet incomplete) contextual generalization, as is expected of grammaticizing constructions.

Table 6.52. Distribution of verb class in PF and temporal SF by data set

<table>
<thead>
<tr>
<th>Verb class</th>
<th>13</th>
<th>15</th>
<th>17</th>
<th>19</th>
<th>20w</th>
<th>20s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic (non-motion)</td>
<td>100</td>
<td>53</td>
<td>67</td>
<td>48</td>
<td>64</td>
<td>44</td>
</tr>
<tr>
<td>Static</td>
<td>--</td>
<td>27</td>
<td>--</td>
<td>30</td>
<td>5</td>
<td>34</td>
</tr>
<tr>
<td>Motion</td>
<td>--</td>
<td>7</td>
<td>17</td>
<td>6</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>Ir</td>
<td>--</td>
<td>3</td>
<td>--</td>
<td>3</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td>Perception</td>
<td>--</td>
<td>4</td>
<td>17</td>
<td>6</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Psychological</td>
<td>--</td>
<td>6</td>
<td>--</td>
<td>7</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>&lt;1</td>
<td>99</td>
<td>2</td>
<td>98</td>
<td>4</td>
<td>96</td>
</tr>
</tbody>
</table>

While this generalization of PF to stative verbs took place, however, the stative verbs in SF were not as stable as Table 6.52 would lead us to believe. While they stood their ground as futures in SF, stative verbs also began being used disproportionately in non-temporal epistemic SF contexts.
Figure 6.7. Proportion of stative verbs in temporal PF and temporal SF

Figure 6.8. Proportion of epistemic SF in all SF, compared to proportion of epistemic SF in all occurrences of stative verbs in SF

Figure 6.7 shows the percentage of stative verbs in temporal SF and temporal PF, a visual presentation of the data in Table 6.52 just discussed. The top line represents the
proportion of temporal SF that occurs with stative verbs. The lower line shows the proportion of temporal PF that occurs with stative verbs: while the SF proportion of stative verbs has remained relatively stable, the proportion of PF with stative verbs has risen from 0% to 20%, at 5% of the PF occurrences in the 17th century, 8% in the 19th and 20th centuries in written data, and 20% in 20th-century speech. Despite this increase, the rate of stative verbs in PF has not yet reached the long-standing rate of about 30% in the temporal SF.

This Figure 6.7—which depicts only future uses of the two forms—might lead us to believe that the encroachment merely involves PF contextual generalization into stative contexts. However, this may be only part of the story. While stative use within temporal SF remained relatively constant, this linguistic context has been losing its status as one of the best contexts for temporal SF. As the PF becomes more common for prediction, the SF becomes more associated with an epistemic reading, and—what is important here—this association of SF with epistemicity is strongest with stative verbs. Figure 6.8 shows the proportion of epistemic SF uses of all SF tokens as a whole and of stative SF tokens. First, as we have previously seen (Figure 5.6), epistemic SF use in general made up only 2% of the SF in the 19th century, 10% in 20th-century plays and 24% in 20th-century speech. In stative verbs, the epistemic SF also increases in frequency, but at an accelerated rate. Since the 19th century, about a quarter (23-25%) of all SF statives have occurred outside the realm of future temporal expression. In 20th-century speech, this rate approaches half, at 41%. The comparative numbers in Figure 6.8, then, show an increasingly strong tendency for speakers to use SF epistemically in
stative contexts, an association that increased disproportionally as the use of the epistemic SF grew.

This result is particularly interesting when we observe the slower yet parallel climb of the PF in stative contexts, which suggests a possible connection between these phenomena. The nature of this connection is not clear. Perhaps as the PF generalized during grammaticization, it weakened the association of future states with the SF construction, allowing the epistemic meanings so easily inferred in stative future contexts to become more salient to speakers, until they were conventionalized as the most frequent context in which the SF was used with epistemic meaning. Or perhaps the ever-frequent stative verbs in SF began to be increasingly associated with epistemicity as the SF lost its ability to express intention (Section 3.5), which allowed the PF to emerge in this context. The quantitative evidence in Figure 6.8 seems to offer support in favor of the second scenario, since the 19th-century rise in in stative epistemic contexts is more impressive than the concurrent rise in stative PF contexts. Whatever the case may be, this is the first piece of what may be thought of as "circumstantial evidence" that the coordinated rise in epistemic SF and changes in PF and SF conditioning is not coincidental.

6.3.2 Interrogatives

The second piece of evidence for the impact epistemic SF development may have on SF future use and therefore variation patterns involving PF is found in one of the constraints on PF-SF variation that reaches significance for the first time in the 20th-century data
sets: sentence type. Table 6.53 shows the rates of interrogative clauses in temporal PF, temporal SF, and epistemic SF by century.117

Table 6.53. Proportion of interrogatives with PF, temporal SF and epistemic SF by data set

<table>
<thead>
<tr>
<th>Century</th>
<th>PF % (N)</th>
<th>Temporal SF % (N)</th>
<th>Epistemic SF % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>0 (0/1)</td>
<td>9 (92/1058)</td>
<td>50 (2/4)</td>
</tr>
<tr>
<td>15</td>
<td>25 (3/12)</td>
<td>11 (83/770)</td>
<td>6 (1/16)</td>
</tr>
<tr>
<td>17</td>
<td>10 (6/59)</td>
<td>7 (38/1270)</td>
<td>32 (7/22)</td>
</tr>
<tr>
<td>19</td>
<td>16 (12/75)</td>
<td>16 (70/432)</td>
<td>40 (17/43)</td>
</tr>
<tr>
<td>20w</td>
<td>18 (14/79)</td>
<td>6 (12/212)</td>
<td>8 (2/120)</td>
</tr>
<tr>
<td>20s</td>
<td>15 (118/768)</td>
<td>4 (15/379)</td>
<td>17 (20/229)</td>
</tr>
<tr>
<td>Total % interrogative</td>
<td>15</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Total N =</td>
<td>153/994</td>
<td>360/4121</td>
<td>49/434</td>
</tr>
</tbody>
</table>

As we can see here, the epistemic SF and the PF have in common the fact that the proportion of interrogative clauses has generally been greater than in the temporal SF.

This tendency is particularly marked for the epistemic SF, which accounts for 12% of interrogatives and only 5% of declaratives, a significant difference ($p < .000$, $x^2 = 42.65468$) seen in Table 3.24, repeated below as Table Table 6.54, which combines data from all centuries. Returning to Table 6.53, it is in the 20th century that we see the lowest rates for interrogatives in temporal SF, with 6% (12/212) in written texts and 4% (15/379) in spoken data (bottom two rows of Table 6.53), significantly lower than the 19th-century interrogative rates with temporal SF ($p \leq .0000$, $x^2 = 39.36498$). In the latter data set, the PF and epistemic SF display relatively similar interrogative rates of 15%

117 Percentages represented here include only the proportion of each use (PF, temporal SF, and epistemic SF) that occurred in interrogative clauses. In interest of readability, this table does not include percentages or Ns for declarative clauses, which can be calculated from the numbers provided; thus, percentages in this table do not add up to 100.
(153/994) and 17% (49/343), respectively, a difference that is significant only at the .05 level ($p \leq .0408$, $x^2 = 4.185627$).

Table 6.54. Distribution of temporal and epistemic SF by sentence type, all data sets

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Temporal</th>
<th>Epistemic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N  %</td>
<td>N  %</td>
<td>N  %</td>
</tr>
<tr>
<td>Interrogative</td>
<td>354 88</td>
<td>49 12</td>
<td>403 100</td>
</tr>
<tr>
<td>Declarative</td>
<td>3781 95</td>
<td>180 5</td>
<td>3961 100</td>
</tr>
</tbody>
</table>

$p < .0000, x^2 = 42.65468$

With sentence type, then, as with verb class, we see a reflection of epistemic SF distributional patterns in the realm of 20th-century PF-SF variation in future expression. As the epistemic SF rises in frequency, the PF comes to be (significantly) favored in interrogative contexts. Meanwhile, rates of interrogatives in both epistemic SF and the PF show no particular change; instead, it appears that the SF is used ever less in interrogative future contexts.

Before I discuss the possible reasons for the decrease of interrogatives with temporal SF, I would like to address the essential question of function. If PF and epistemic SF have similar rates of interrogative clauses, does that mean that speakers use these forms in interrogatives to perform the same communicative tasks? In terms of the semantics of conjecture, such a proposal makes little sense, since the PF is associated with certainty and the epistemic SF with uncertainty (Section 5.3.5.1). The similarity of PF and epistemic SF can also not be explained in terms of clause type, since PF is favored in subordinate clauses (Table 5.44) and the epistemic SF occurs more often than the temporal SF in main clauses (Table 3.21). I would argue that the similar rates of interrogatives with these two form-meaning pairings have no direct relationship with each other, i.e., it is not simply because they are both associated with a particular meaning or syntactic feature.
To understand why both the PF and epistemic SF have been favored in interrogatives since their incipient stages (see Table 5.42, Table 3.23 and Figure 6.10), we must consider again the notions of intention and prediction. We will see that speakers do not use these forms in interrogatives with the same motivations at all: epistemic SF is highly constrained by its inability to express intention, while the PF is associated in earlier centuries with intention and subsequently with prediction. As we saw in Section 5.3.5.1, the PF's ability to express intention—as measured through the co-occurrence of first-person singular subjects—rose in the 19th century, and then gradually declined. On the other hand, the temporal SF's ability to express intention was declining between the 17th and 20th centuries. We also saw in Section 3.4.2 that the epistemic SF's inability to express intention is nearly categorical; by definition, the semantic feature of intention is bleached in epistemic grams originating in futures.

If one function of interrogatives is to ask about intentions, as surely it must be, then we would expect the PF and the SF to be used differently. Interrogatives regarding intention are most likely to appear in the second-person singular, as in (6.153) (we probably inquire about the intentions of our interlocutors more than those of third parties); we would thus hypothesize that the PF would show an increase in second-person uses, that the temporal SF would show a decrease, and that the epistemic SF would show generally low rates of co-occurrence with second-person interrogatives.

(6.169) MATILDE. - (Queriendo detenerla.) ¿Qué **has a hacer**?...repara... (Conjuración, Act III, Scene I, 19c.)
'MATILDE: [Wanting to stop her.] What are you going to do (PF)?... wait...'

En fin, ¿me **dirás** qué tienes? En un instante te encuentro... qué se yo... (El afán, Act IV, Scene IV, p. 283, 19c.)
'Finally, will you tell me (SF) what's wrong? In just an instant I find you... what do I know...'
Likewise, first-person interrogatives make little sense with intention meaning.

Inquiring about intention, of course, is not the only function of interrogatives with future forms. As noted in the case of PF in particular (and SF prior to the 20th century), interrogatives in conjunction with future forms can express a special meaning of rhetorical incredulity, which, I have argued, is a prediction meaning in a specialized context (Section 4.3.2). In fact, it is in this specialized context that we find first-person interrogatives, as in (6.170).

(6.170)"-Pero usted está loco! ¿Cómo le voy a echarte esa cosa roja ahí en mitad del grano, mitad de la cara?" (COREC, CACON022C, 20s)
"But you're crazy! How am I going to put (PF) that red thing on half of the pimple, half of my face?"

In (6.170), the speaker expresses disbelief about the prediction that she will put a patch on her face to clear up her pimple. Given the use of first-person singular in future interrogatives to express rhetorical incredulity, and also given the finding that the PF is the only form associated with this context in the 20th century (Section 4.3.2), we may expect to find that the PF is the preferred form in first-person interrogatives, at least in the 20th century. I have suggested, then, three associations regarding grammatical person in interrogative sentences, which I hope will explain why both the PF and epistemic SF occur more often in interrogatives than the temporal SF: i) intention with second person, and ii) prediction (particularly the rhetorical incredulity context) with first person singular; iii) interrogative contexts are ideal for emphasizing epistemic modality because some interrogatives themselves express this same modality (Narrog 2005:679-680).

In Table 6.55, Table 6.56, Table 6.57, and Table 6.58, we see the distribution of first and second-person subjects in interrogative clauses with PF, temporal SF and
epistemic SF in four data sets. In the 17th century, seen in Table 6.55, second-person (i.e. likely intention) uses are associated most strongly with the PF, accounting for 67% (4/6), compared to 16% (18/88) in temporal SF and 14% (1/7) in epistemic SF. The only form in this data set that is used with first-person interrogatives is the temporal SF, with 24% of its occurrences in first person singular. The epistemic SF occurs 86% (6/7) of the time in interrogatives with neither first nor second person subjects. This suggests that interrogatives with the PF in the 17th century were mainly used to question intention, while in the temporal SF prediction was the main function.

Table 6.55. Distribution of subject within interrogative clauses, 17th century

<table>
<thead>
<tr>
<th></th>
<th>PF % (N)</th>
<th>Temp SF % (N)</th>
<th>Epistemic SF % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s</td>
<td>-- (0)</td>
<td>24 (27)</td>
<td>-- (0)</td>
</tr>
<tr>
<td>2s</td>
<td>67 (4)</td>
<td>16 (18)</td>
<td>14 (1)</td>
</tr>
<tr>
<td>Other</td>
<td>33 (2)</td>
<td>48 (54)</td>
<td>86 (6)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total N</td>
<td>6</td>
<td>88</td>
<td>7</td>
</tr>
</tbody>
</table>

In the 19th century, in Table 6.56, the PF is no longer so strongly associated with intention meaning: 25% (3/12) of PF interrogatives occur with first-person singular, and 42% (5/12) with second-person singular. At the same time, temporal SF is used only 8% (6/70) of the time in second-person with interrogatives, and epistemic SF does not occur at all with this person in interrogatives (0/17). This suggests that in the 19th century, the PF was used in interrogatives with both intention and prediction meaning, while intention meaning was minimal in interrogative uses of both epistemic and temporal SF. These results also suggest the transfer of rhetorical incredulity uses from temporal SF to PF during this period, as the relative frequency of first-person interrogatives increased with PF to 25% (3/12) and with SF it decreased to 14% (10/70).
Table 6.56. Distribution of subject within interrogative clauses, 19th century

<table>
<thead>
<tr>
<th></th>
<th>PF % (N)</th>
<th>Temp SF % (N)</th>
<th>Epistemic SF % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s</td>
<td>25 (3)</td>
<td>14 (10)</td>
<td>6 (1)</td>
</tr>
<tr>
<td>2s</td>
<td>42 (5)</td>
<td>8 (6)</td>
<td>-- (0)</td>
</tr>
<tr>
<td>Other</td>
<td>33 (4)</td>
<td>77 (54)</td>
<td>94 (16)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total N =</td>
<td>12</td>
<td>70</td>
<td>17</td>
</tr>
</tbody>
</table>

The 20th century data sets show a similar picture. In both written and spoken data sets, shown in Table 6.57 and Table 6.58, the PF is the only use that occurs with first-person interrogatives, at 14% (2/14 and 16/118, respectively); there are no first-person interrogative occurrences in the 20th-century data. This shows that the temporal SF has lost any association with rhetorical incredulity uses. The temporal SF also appears to have lost nearly all association with intention meaning, a meaning that epistemic SF never had. In the case of temporal SF, second-person subjects make up 8% (1/12) of interrogatives in written data and 7% (1/15) in spoken data. In the case of epistemic SF, there is only one occurrence of a second-person interrogative, shown in (6.171), which is a prediction use in a message left on an answering machine.

(6.171) Vale Ritchie, pues - hasta mañana entonces. ¿Dónde estarás golfo? (COREC, CPCON006A, 20s)
'Okay, Ritchie, well – until tomorrow then. Where could you be (SF) vagabond?'

It appears that in the 20th century, then, the PF was the sole future form used for both rhetorical incredulity and intention in interrogatives, which may explain why the SF decreased in frequency during this century. Speakers began to find the SF simply less useful.
Table 6.57. Distribution of subject within interrogative clauses, 20th-century writing

<table>
<thead>
<tr>
<th></th>
<th>PF % (N)</th>
<th>Temp SF % (N)</th>
<th>Epistemic SF % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s</td>
<td>14 (2)</td>
<td>-- (0)</td>
<td>-- (0)</td>
</tr>
<tr>
<td>2s</td>
<td>50 (7)</td>
<td>8 (1)</td>
<td>-- (0)</td>
</tr>
<tr>
<td>Other</td>
<td>36 (5)</td>
<td>92 (11)</td>
<td>100 (2)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total N =</td>
<td>14</td>
<td>12</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 6.58. Distribution of subject within interrogative clauses, 20th-century speech

<table>
<thead>
<tr>
<th></th>
<th>PF % (N)</th>
<th>Temp SF % (N)</th>
<th>Epistemic SF % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s</td>
<td>14 (16)</td>
<td>-- (0)</td>
<td>-- (0)</td>
</tr>
<tr>
<td>2s</td>
<td>35 (41)</td>
<td>7 (1)</td>
<td>5 (1)</td>
</tr>
<tr>
<td>Other</td>
<td>52 (61)</td>
<td>93 (14)</td>
<td>95 (19)</td>
</tr>
<tr>
<td>Total %</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total N =</td>
<td>118</td>
<td>15</td>
<td>20</td>
</tr>
</tbody>
</table>

If the SF was losing functions as the PF was gaining them, then why would 17% of epistemic SF uses in 20th-century speech occur in interrogatives (Table 6.53)? It certainly is not because of an elevated ability to express intention, since first-person co-occurrence with epistemic SF is extremely rare. It may perhaps be that interrogatives correlate with epistemicity. Narrog (2005:679) defines modality as "a linguistic category referring to the factual status of a state of affairs. The expression of a state of affairs is modalized if it is marked for being undetermined with respect to its factual status, i.e., is neither positively nor negatively factual." We can see the meaning of epistemic SF interrogatives as depicting unknown factual status in the epistemic SF interrogatives in subordinate clauses, 100% (6/6) of which have the head verb no saber 'not know'.

Here, then, we see that the rise of epistemic uses of SF, especially in the semantically harmonic context of interrogatives, may have led to an intensification of the
association of interrogatives in future tense with the PF, which was always favored over SF due to its greater ability to express intention, and thus applicability to a wider range of contexts. Therefore, while there was not a direct push-pull relationship between epistemic SF and the PF, the emergence of the former into the grammar intensified previously insignificant tendencies in PF-SF variation, leading to a change in the division of labor in future expression in the 20th century.
6.3.3 Temporal adverbials

In the case of both stative verbs and interrogatives, the results seem to suggest a correlation between the generalization of the PF into a particular linguistic context within the future realm and the contexts in which epistemic SF tends to occur. As the epistemic SF rises in frequency, the PF comes to express futurity in the linguistic contexts in which the SF is most likely to be used non-temporally. Another constraint that comes up as significant in 20th-century PF-SF variation is that of the presence of a temporal adverbial. This constraint was also operative in the 17th century, and displayed the same tendencies: PF is most favored in the absence of a temporal adverbial, less favored by a specific temporal adverbial, and least favored by a non-specific temporal adverbial. It then was not selected as significant in the 19th century, but the direction of effect remained the same. Temporal adverbial presence returned as one of the weaker constraints in the 20th-century texts, and became the constraint with the highest magnitude of effect in 20th-century speech (see Chapter 5). The SF is favored with nonspecific adverbials, as in (6.172)-(6.173), while the PF is favored when no adverbial is present, or if there is an adverbial, when it is specific, as in (6.174)-(6.175).

(6.172) Que tarde o temprano se *va a venir* a España (COREC, CBCON014A, 20s) 
'Sooner or later he's going to come (PF) to Spain'

(6.173) Y *aprenderé* alemán algún día. (COREC, CACON012A, 20s) 
'And I will learn (SF) German someday'

(6.174) Oye, mañana - hoy es - martes, mañana por la mañana te *llamaré* a ver si me puedes hacer un favor. (COREC, CPCON006A, 20s) 
'Listen, tomorrow – today is – Tuesday, tomorrow in the morning I'll call (SF) you to see if you can do me a favor'

(6.175) está la tía metida en casa y tal y - nos *vamos a ir* mañana a verla. (COREC, CACON019B, 20s) 
'the girl is stuck at home and such and – we're going to go (PF) see her tomorrow'
In my opinion, there are three plausible ways to interpret this result. The first, and most obvious, is that this constraint represents a continuation of the same constraint found in the 17th century. While this must indeed be the case, such an analysis leaves unanswered why this constraint would appear to weaken in the 19th century, only to rise again in the 20th, becoming the single most important factor in PF-SF variation in 20th-century speech (see Section 4.3.4). Figure 6.9 shows the rate of occurrence of both specific and non-specific temporal adverbials with temporal uses of the PF and the SF since the 13th century.
Regarding general tendencies, Figure 6.9 shows that the temporal SF has always occurred more often with temporal adverbials than the PF. These general patterns, while certainly of interest, are not what catches our attention, however. What is impressive in Figure 6.9 is the dramatic rise in SF co-occurrence with nonspecific temporal adverbials, from 10% in the 19th century, to 20% in 20th-century texts and 28% in 20th-century speech. No other context shows such a sharp increase.

The increase in the rate of co-occurring temporal adverbials with temporal SF, specific and nonspecific combined, from 18% in the 17th century data to 40% in the 20th century speech data (Figure 6.9), is due to the development of the non-temporal epistemic use of the SF, whose frequency relative to temporal SF use rises from 2% to 24% in the same period (Figure 3.4), which would threaten speakers' (and crucially hearers') association of the SF with future temporal reference. Hearers play an important role in that it may be with the hearer's state of knowledge in mind that a speaker provides extra information to ensure understanding. In other words, if a speaker wants to be understood
as referring to a future moment, and the same speaker construes her hearer's associations of SF to include contexts outside futurity, she may produce a temporal adverbial alongside the SF. Thus, as speakers begin to use—and hear—the SF more frequently in non-future uses, the meaning of SF no longer includes futurity as an obligatory component.

For example, in (6.176), the speaker is making a prediction about the behavior of her students at the next class she is to teach, on Wednesday: she is not sure if they will bring the balls that day. She signals twice that she is unsure about this prediction, with no sé si 'I don't know if'. As noted previously, such contexts of uncertainty in prediction are those in which the SF is most likely to occur in non-future contexts. At the same, the speaker in (6.176) is referring to other, habitual behavior of her students, that they always forget the balls (se les olvidan 'they forget them'). Since SF semantics no longer obligatorily include future temporal reference, both times the speaker uses SF traerán 'they will bring', she uses a temporal adverbial as well, ese día 'that day' and el miércoles 'on Wednesday', to signal future temporal reference.

(6.176)- Y, pero, entonces la clase es de seis a siete.
- De seis a siete.
  - Pero, claro, yo tampoco te quiero a ti molestar o - eso.
  - Que no, que no te preocupes. Ya verás.
  - Bueno, yo voy a las seis y si algún rato pues eso, pues me siento y me callo o -
  - Bueno, sí, además ese día, no sé si traerán unas pelotas porque les iba a enseñar a darse unos masajes. ¿Sabes?
  - Ah.
  - Entonces pero no sé si se les ha - olvidará o qué porque llevamos con las pelotas desde no sé cuándo, ¿sabes? y - y se les olvidan, entonces no sé si el miércoles traerán las pelotas y - si no, vamos a hacer canciones y juegos por ser el último día, ¿sabes? (COREC, CCCON029C, 20s)

- And, but, then the class is from six to seven.
- From six to seven.
- But, of course, I don't want to bother you, either, or – yeah.
- (That) no, (that) don't worry. You'll see.
- Well, I'll go at six and if for a while well yeah, well I'll sit and be quiet or –
- Well, yeah, also that day. I don't know if they'll bring (SF) some balls because I was going to teach them to give massages. You know?
- Ah.
- So but I don't know if they have – they'll forget or what because we've been talking about the balls since I don't know when, you know?, and – and they forget them, so I don't know if on Wednesday they'll bring (SF) the balls and – if not, we're going to do songs and games because it's the last day, you know?'

In (6.177), the temporal adverbial *dentro de un año o dos* 'within a year or two' is necessary to the meaning of the utterance; without this adverbial, the utterance could easily be understood as referring to the present, as in *yo creo que no estará terminada* 'I think it's probably not done yet'.

(6.177) Porque tienen un lio con el Canal de Isabel Segunda y con no sé qué sitio más - y bueno, un lio. Yo creo que hasta *dentro de un año o dos* no *estará* terminada. (COREC, CACON021A, 20s)
'Because they have a mess with the Canal of Isabelle the Second and with I don't know what other place – and well, a mess. I think that even within a year or two it won't be (SF) done.'

Thus, the SF begins to require contextual support in some occurrences in order to effectively communicate future temporal reference. The contextual support that would best ensure the temporal reading of this newly polysemous form is a temporal adverbial.

In this case, it appears that the types of adverbials most often chosen were those with which the SF had always (since the earliest Old Spanish data) been most likely to occur: nonspecific adverbials. It is without doubt that this dramatic rise in SF co-occurrence with nonspecific temporal adverbials is behind the return of adverbial presence as a significant factor in 20th-century PF-SF variation. This is yet another link—albeit indirect—between the rise in epistemic SF and the shifts in PF-SF variation patterns: the divergence of SF—and loss of default status—led to an increase of temporal adverbials
with temporal SF, and the increase in the rate of co-occurring temporal adverbials led to a restructuring in the ordering of factors conditioning 20\textsuperscript{th}-century PF-SF variation.

6.4 Overall tendencies: grammaticization, variation and divergence

In Section 6.3, I discussed three contextual factors, verb class, sentence type, and temporal adverbials, in which the patterns found in 20\textsuperscript{th}-century PF and the rise and distributional tendencies of the epistemic SF appear to be connected. It is important to remember that the epistemic SF is a temporal SF bleached of future temporal reference and intention meaning, i.e., it is a more grammaticized manifestation of the same form (see Chapter 3; cf. Bybee, Pagliuca & Perkins 1991). Since this is the case, and categories are inherently gradient (cf. Thompson 2002; Company 2003), we may want to adopt the methodology proposed by Schwenter and Torres Cacoullos (2005) (cf. Torres Cacoullos 2001): include competing forms' entire paths of grammaticization in the variable rule analyses. In their study of Preterite-Present Perfect variation in present-day Peninsular Spanish, the forms displayed a set of functions along the perfect to perfective grammaticization path, allowing for a function-based study that took into account entire paths of grammaticization. Is a broad function-based analysis like Schwenter and Torres Cacoullos' (2005) possible in the case of 20\textsuperscript{th} c. PF-SF variation, since these two constructions' grammaticization paths do not overlap completely in terms of function; (as we have seen, the PF is not used as a non-future epistemic marker in these Peninsular Spanish data)? That is, is this methodology, which has proven valuable in tracing a form's meaning as it shifted within the same category, similarly useful when grammaticization stretches across categories?
In order to test whether the model of an envelope of variation based on
diachronically-related functions in grammaticization (Torres Cacoullos 2001, Schwenter
and Torres Cacoullos' (2005) can be appropriately extended to cases like that found with
the SF (in which there is not complete functional overlap of paths of grammaticization
between variants), I have returned to the 20th-century spoken data, where the epistemic
SF is most frequent. I have then taken all occurrences of both PF and SF, regardless of
meaning (excluding fixed phrases), and done a form-based analysis of the variation
between these two forms. The results of this form-based analysis, side-by-side with the
results from the function-based analysis, i.e. of future temporal reference uses only,
originally presented in Chapter 5, are shown in Table 6.59. If the inclusion of the entire
grammaticization path is useful even when there is not complete (diachronic) functional
overlap, we would expect to find that the form-based VARBRUL analyses would offer
detailed information about the SF as a form that was superior to the information gleaned
from function-based comparisons.
Table 6.59. Variable rule analyses of the contribution of factors selected as significant to occurrence of PF; 20\textsuperscript{th}-century spoken data

Form: Input probability: 0.622 (61%), N = 775/1274
Function: Input probability: 0.692 (67%), N = 768/1147

<table>
<thead>
<tr>
<th>Factor Group</th>
<th>Form % PF</th>
<th>Probability</th>
<th>Function % PF</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temporal specificity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>66</td>
<td>.56</td>
<td>73</td>
<td>.57</td>
</tr>
<tr>
<td>Specific</td>
<td>60</td>
<td>.47</td>
<td>59</td>
<td>.42</td>
</tr>
<tr>
<td>Nonspecific</td>
<td>36</td>
<td>.25</td>
<td>37</td>
<td>.22</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Verb class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic (non-motion)</td>
<td>73</td>
<td>.63</td>
<td>74</td>
<td>.58</td>
</tr>
<tr>
<td>Motion (incl. \textit{ir})</td>
<td>67</td>
<td>.56</td>
<td>69</td>
<td>.53</td>
</tr>
<tr>
<td>Stative/percep./psych.</td>
<td>44</td>
<td>.32</td>
<td>55</td>
<td>.36</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sentence type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interrogative</td>
<td>77</td>
<td>.67</td>
<td>88</td>
<td>.78</td>
</tr>
<tr>
<td>Declarative</td>
<td>59</td>
<td>.48</td>
<td>64</td>
<td>.46</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Polarity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>67</td>
<td>.60</td>
<td>74</td>
<td>.61</td>
</tr>
<tr>
<td>Affirmative</td>
<td>60</td>
<td>.49</td>
<td>66</td>
<td>.49</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clause type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subordinate</td>
<td>65</td>
<td>[.53]*</td>
<td>70</td>
<td>[.52]</td>
</tr>
<tr>
<td>Main</td>
<td>59</td>
<td>[.49]</td>
<td>65</td>
<td>[.49]</td>
</tr>
</tbody>
</table>

Form: Log likelihood = -762.351; \( p = .028 \); Chi-square/cell = 1.4111
Function: Log likelihood = -648.430; \( p = .039 \); Chi-square/cell = 1.1124
*Square brackets [ ] indicate that this effect does not achieve statistical significance.

As we see in Table 6.59, a form-based analysis of PF-SF variation in 20\textsuperscript{th}-century speech reveals about the same story as a function-based analysis. Again, temporal specificity, verb class, sentence type and polarity are selected as significant, and the direction of effect is the same in all cases. Given this similarity in results, it appears that including occurrences outside the traditional scope of the variable context has at least not changed the basic analysis in any way. The only difference, in fact, is the ordering of the factor
groups by magnitude of effect, shown in boxes; in the form-based analysis, temporal specificity and verb class share top rank, both with a range of 31, followed by sentence type with a range of 19, and finally polarity, with a range of 11.

Thus, including uses of SF in non-future contexts in the variable rule analysis obscures the impact of interrogative favoring the PF, as this factor group is demoted from second-most important in the function-based analysis to third-most important in the form-based analysis. This is due to the fact that, while temporal SF has a low rate of occurrence with interrogatives, epistemic SF has a higher rate, similar to that of the PF. By combining temporal and epistemic SF, the clear picture of a characteristically lower rate of interrogatives with temporal SF is muddled.

These results suggest that variable rule analyses need to be carried out only on function-based variants, though, of course, considering all uses of forms is important (see Sections 6.2-6.3 on the insights gleaned from quantitative investigation of epistemic use of SF). In the case of the SF form, the differences between temporal SF and epistemic SF are, in fact, statistically significant (Section 3.5). There are three contexts in which we find significant differences between temporal SF and epistemic SF: temporal specificity, verb class and sentence type, shown in Tables 6.60, 6.61 and 6.62, which repeat information in Table 3.23 on sentence type, present in a different way information in Table 3.7 and Table 3.12 on verb class, and present new information on temporal specificity.

118 Chapters 3 and 4 are also evidence of the usefulness and necessity of form-based studies as a preliminary step in any analysis of variation, since it is through such analyses that one can determine the envelope of variation.
Table 6.60. Relative frequency of temporal and epistemic SF according to temporal specificity, all data sets

<table>
<thead>
<tr>
<th>Temporal specificity</th>
<th>Temporal N</th>
<th>%</th>
<th>Epistemic N</th>
<th>%</th>
<th>Total N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not present</td>
<td>3342</td>
<td>94</td>
<td>212</td>
<td>6</td>
<td>3554</td>
<td>100</td>
</tr>
<tr>
<td>Present</td>
<td>779</td>
<td>98</td>
<td>17</td>
<td>2</td>
<td>796</td>
<td>100</td>
</tr>
</tbody>
</table>

$p < .0000, x^2 = 19.12274$

Table 6.61. Relative frequency of temporal and epistemic SF according to stativity, all data sets

<table>
<thead>
<tr>
<th>Stativity</th>
<th>Temporal N</th>
<th>%</th>
<th>Epistemic N</th>
<th>%</th>
<th>Total N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stative</td>
<td>1264</td>
<td>88</td>
<td>179</td>
<td>12</td>
<td>1443</td>
<td>100</td>
</tr>
<tr>
<td>Other</td>
<td>2869</td>
<td>98</td>
<td>50</td>
<td>2</td>
<td>2919</td>
<td>100</td>
</tr>
</tbody>
</table>

$p < .0000, x^2 = 221.915$

Table 6.62. Relative frequency of temporal and epistemic SF according to sentence type, all data sets

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Temporal N</th>
<th>%</th>
<th>Epistemic N</th>
<th>%</th>
<th>Total N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interrogative</td>
<td>354</td>
<td>88</td>
<td>49</td>
<td>12</td>
<td>403</td>
<td>100</td>
</tr>
<tr>
<td>Declarative</td>
<td>3781</td>
<td>95</td>
<td>180</td>
<td>5</td>
<td>3961</td>
<td>100</td>
</tr>
</tbody>
</table>

$p < .0000, x^2 = 42.65468$

The existence of significant differences, not surprising given the definition of exaptation offered by Company (2003), suggests that, when a grammaticization path crosses a category boundary, it may no longer be appropriate to include all occurrences along the path in the envelope of variation (variable context) of a variationist study. In this case, the contexts that favor a temporal reading of SF and the contexts that favor an epistemic reading of SF have diverged in some contexts (interrogatives, statives, temporal adverbs) and moved closer in others (polarity).

6.5 Summary of results and implications: shifting defaults

This chapter has presented a picture of two forms, the PF and the SF, grammaticizing side-by-side, pulling on each other and shaping the results of quantitative analyses. I have
<table>
<thead>
<tr>
<th>Task</th>
<th>Condition</th>
<th>Time</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task A</td>
<td>Condition 1</td>
<td>10</td>
<td>85%</td>
</tr>
<tr>
<td>Task B</td>
<td>Condition 2</td>
<td>15</td>
<td>70%</td>
</tr>
</tbody>
</table>

*Note: Performance values are based on a set of standard conditions.*
suggested that the particular manifestations of the process of grammaticization of the conservative form has, in this case, played an active role in the negotiation of constraints on variation with the innovative form. This is not meant to imply, however, that the SF's path was able to influence the path of the PF; instead, the SF's continued grammaticization affected the way it was used as a future, which, in turn, affected the way speakers chose one form over the other. Without an understanding of how the SF was affected by further grammaticization in the 20th century, we would be hard-pressed to explain the shifts in conditioning on variant choice.

In the case of the PF, we see a form that has undergone contextual generalization as it has taken on functions it did not tend to perform in its lexical beginnings, most especially prediction; retention of PF's lexical origins appears in the 20th century only in stative contexts, the last stronghold of the SF as a future form. In the 20th century, however, we see changes in the conditioning of PF-SF variation that do not seem to be related to PF generalization, nor are they retentions from SF, such as obligation meaning. Instead, these changes involve the strengthening of constraints that appeared to be weakening in the 19th century. I have argued here that this strengthening is related to the continued grammaticization of the SF in the 20th century, which included exaptation and a weakened ability to express futurity. The evidence for this is found in significant changes in the rates of co-occurring contextual features in 20th-century SF, which are both a result of SF divergence and the reason behind shifts in the division of labor in 20th-century Peninsular Spanish future expression.

There was one general change and three specific changes examined here: i) a diachronically parallel rise in the relative frequencies of the PF and of epistemic SF; ii)
the generalization of PF into stative contexts as speakers began to increasingly associate stative SF with epistemic meaning; iii) a decrease in interrogatives with temporal SF as SF interrogatives became increasingly associated with epistemic use and the SF simultaneously lost its ability to express intention; and iv) a sharp increase in the co-occurrence of temporal adverbials with temporal SF as the SF began to be bleached of temporal meaning in its epistemic uses. The cases of interrogatives and temporal adverbials lead me to conclude that, for speakers, the unmarked SF is no longer a future form perhaps embued with epistemicity, but rather both an epistemic marker and a future marker.

The SF may correctly be analyzed as having two grammatically distinct functions, one temporal, one not. The significant differences in distributional patterns between the temporal and epistemic SF (see Chapter 3) suggest that it may be inappropriate to include the entire path of grammaticization in variable rule analyses in all cases, since such a move can obscure the complex dynamics of divergence in a construction. Caution in broadening the variable context to include tokens outside the category of the functional space in question does not imply that any piece of the path of grammaticization can be ignored. In quantitative study, sensitivity to uses such as the epistemic SF may be the key to understanding otherwise inexplicable changes in distribution within the variable context. How, for example, would we explain the sudden 20th-century rise in temporal adverbials with the temporal SF if we were to ignore the simultaneous rise in epistemic SF? We could not. In the case of exaptation, in which there is a category change in the process of grammaticization—in this case from tense marker to mood marker—the conservative forms used within the "old" category (e.g. tense) may diverge from the
newer use in what could be construed as a last-minute ditch to salvage their own identity as speakers use a growing number of contextual clues to support the bleaching meaning of (functionally) moribund forms. In this case, then we would see an SF that has temporal readings more often in the contexts that are less favorable to the epistemic SF; such a move would make it appear that the SF's competitor—the PF—was moving in parallel step with the epistemic SF. This pattern is found with the increasing co-occurrence of temporal adverbials with the temporal SF and the decline of interrogatives with temporal SF. Figure 6.10 portrays these tendencies.

Figure 6.10. Direction of effect for overt temporal specification/modification, interrogatives and stative verbs on temporal SF, epistemic SF and PF

<table>
<thead>
<tr>
<th>Feature</th>
<th>PF</th>
<th>Temp SF</th>
<th>Epistemic SF</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp. specificity</td>
<td>disfavored</td>
<td>favored</td>
<td>less likely</td>
<td>Temp ≠ Epist</td>
</tr>
<tr>
<td>Interrogatives</td>
<td>favored</td>
<td>disfavored</td>
<td>more likely</td>
<td>Temp ≠ Epist</td>
</tr>
<tr>
<td>Stative verbs</td>
<td>~disfavored</td>
<td>favored</td>
<td>more likely</td>
<td>Temp ≈ Epist</td>
</tr>
</tbody>
</table>

Given the evidence presented here, it appears that the divergence of a construction in competition can affect its relative frequency in particular contexts (factors), which can, in turn, cause changes in the ordering of conditions on variation. Furthermore, it appears that the semantic bleaching of an older form can trigger further grammaticization and semantic generalization of a newer form into the former's erstwhile semantic territory (see analysis of Figure 6.8 in Section 6.3.1). The contexts into which the newer form generalizes may be directly linked to the contexts into which the older form generalizes. This relationship can create a kind of reflection, in which a new form "pushes" the older form outside the envelope of variation as it encroaches upon its territory.
7 Summary and conclusions

7.1 The approach

This study aimed to take the tool for quantitative analysis of linguistic variation (variationist methodology) and apply it to one of the most exciting theories of language change (grammaticization theory) in the highly volatile domain of Spanish future expression. This approach was chosen due to the complementary nature of these ideas (cf. Torres Cacoullos 2001) in a unique ability to both describe and trace in great detail the changes in constraints on variation and to explain the motivations that may drive such changes. Both theoretical and empirical, a variationist approach to grammaticization offers not only valuable localized information about one process in one functional space in one variety over a specific period of time, but also contributes to our understanding of how language change proceeds and thus, from a functionalist perspective, the very essence of language.

This study focused on two forms in Spanish that have been in viable competition since the 17th century: the Periphrastic Future and the Synthetic Future. The subject of heated debate not only for Spanish, but also for many other languages, future expression is an ideal site for an empirical study of variation over time because change has been drastic and surprisingly fast. In only 300 years, the PF has managed to grow from a very weak contender for future expression in written texts to the preferred form in speech in Peninsular Spanish. The unfolding of this exciting drama has left linguists in search of explanations. Though still incomplete, this study, taking a usage-based approach to language change, is perhaps the most comprehensive quantitative explanation of this much-studied Spanish future to date. It includes both form-based studies of the PF and
SF, as well as a function-based study that zeroes in on the space in which these forms compete.

7.2 The analysis

7.2.1 Form-based studies

Chapters 3 and 4 presented the results of form-based studies of the SF and the PF, respectively. The purpose of these form-based studies was twofold. First, the intention was to examine general rates of occurrence of these forms, as well as identify any changes in rate or distributional tendencies. The second purpose, critical to an analysis of variation, was to determine the various meanings each form could have, and to distinguish different contexts of use (such as "general truth" in SF) from clear divergence (such as epistemic vs. temporal SF).

In Chapter 3, which focused on the SF, I argued that the SF has two meanings: a temporal meaning and a non-temporal epistemic meaning. These meanings, of course, are diachronically related, with the latter constituting a semantically bleached version of the former. I demonstrate that these should, indeed, be considered different meanings of SF through the comparison of the distributional patterns of both uses. This comparison reveals significant differences in tendencies regarding verb class, sentence type, polarity, and clause type. This distinction does not mean, however, that meaning change was abrupt, but rather that the inherent modality of the future (i.e. prediction about an event subsequent to a point of reference) became more salient as the intention meaning of SF bleached.

Regarding other posited meanings of the SF, such as "imperative" and "general truth," I argue that these should rather be seen as constructions that are associated with
the SF, not as separate meanings unique to the SF. This is because the SF continues to contribute its basic future meaning of prediction in these constructions, and because these meanings can occur with other future forms as well. In the case of imperative, interpretation relies upon gauging the intensity of the speaker's assertion. In the case of "general truth" constructions, I argue that it is not the SF, but rather a nonspecific subject, that contributes the generality meaning. The PF has also occurred in both constructions.

In Chapter 4, which looks at the PF, we see that, unlike the temporal SF, which appears relatively stable in terms of distribution, the PF has slowly generalized into more contexts while rising dramatically in frequency. In 20\textsuperscript{th}-century speech, the PF is in fact more frequent than the SF. The PF also shows very few signs of divergence, with only a few occurrences not demonstrating possible temporal interpretations. Nevertheless, the bleaching of the original motion meaning can be seen in a decrease in co-occurrence with locatives. The productivity of the PF in terms of lexical diffusion, as measured by type-token ratio, is shown to increase in written data since the 17\textsuperscript{th} century.

7.2.2 Function-based study

Chapters 3 and 4, then, are portraits of two grammaticizing forms. One is late in its career, dropping in frequency and changing category. The second is young and expanding, taking on new functions and increasing in frequency dramatically. While these portraits tell us a great deal about each of these forms, and bear out the predictions of grammaticization theory, we know that this is not the whole story. These forms have gone about their business side-by-side, perhaps pushing or pulling at each other, each carrying the weight of its own origins. Chapter 4 is dedicated to a diachronic look at approximately 300 years of this interaction.
This diachronic study is carried out through a series of variable rule analyses that take a synchronic snapshot of the conditioning on PF-SF variation in four data sets: the 17th century, the 19th century, 20th-century written texts and 20th-century speech. With regard to the PF, various contextual factors were hypothesized to reflect its lexical origins, including use with dynamic verbs, occurrence with specific temporal adverbials and positive polarity. It was further hypothesized that these constraints on PF-SF variation would diminish in strength as the PF bleached of original allative meaning to become a more generalized future. This was indeed the case; all three of these constraints lost statistical significance or demonstrated reduced magnitude of effect in the 19th century data set.

The 20th century, however, showed a reversal in this weakening process in the case of temporal adverbials, as well as the strengthening of other constraints, which did not have plausible explanation in semantic persistence in PF. Poplack and Malvar (forthcoming) posit that the early (19th-century) favoring of Brazilian Portuguese PF in frequent, less marked contexts such as those without temporal specification "foreshadows its eventual role as default future marker." While the end result in the case of Peninsular Spanish may indeed be the assumption of default status for the PF, here the shifts in the magnitude of effect of the temporal specificity constraint are not directly related to changes in PF status. Instead, they are rooted in changes taking place in the SF during the end of its path of grammaticization: the SF is losing components of meaning, such as intention, and thus becoming contextually more limited. The contexts for 20th-century shift in constraints, then, are to be found outside the realm of future expression, in the final stage of grammaticization of modality-based future forms: epistemicity (Bybee,
Pagliuca & Perkins 1991). This process forms the basis for the analysis presented in Chapter 6.

7.3 Explaining change

Despite the care and detail offered in the form- and function-based studies in Chapters 3-5, some phenomena, particularly shifts in 20th-century constraints on variation, could not be explained with recourse to our knowledge about the grammaticization path of the innovative form, i.e. the PF, or even of the lexical origins of the SF. Instead, I propose in Chapter 6 that we combine form- and function-based analyses in order to take into account the fact that 20th-century SF, i.e. the conservative form, is grammaticizing alongside the PF. We see that the SF suffers exaptation from tense to mood—a potential latent for hundreds of years—and that this divergence of SF provokes changes in the realm of future temporal expression. These changes include apparent divergence of the future uses of SF from the epistemic uses in terms of sentence type and co-occurrence with temporal adverbials. At the same time, the encroachment of PF into the traditionally SF territory of stative verbs occurs, as if to take over the space of SF stative verbs as they are increasingly associated with epistemic meanings.

The results regarding epistemic SF and PF-temporal SF variation suggest that, it may be worthwhile to include high-frequency uses of the forms in variation that fall outside the variable context in quantitative analyses. An attempt to stretch the concept of "variable context" to include epistemic SF—a use that is both categorically dissimilar from future uses of SF and does not overlap in function diachronically with PF—failed, in contrast with the usefulness of including the entire path of grammaticization in the variable context in situations in which functional overlap is (diachronically) complete and
Exaptation is not a concern (Torres Cacoullos 2001; Schwenter & Torres Cacoullos 2005). This failure, however, does not mean that such uses should be ignored; instead, we may wish to examine frequency and distributional patterns of these "other" uses in form- and function-based analyses, as they may hold the key to otherwise inexplicable changes in conditioning. Much of Poplack's pioneering work (e.g. Poplack & Turpin 1999; Poplack & Tagliamonte 2001) aimed to identify differences in conditioning across language varieties or diachronically, and the explanations for these differences were often identified as due to phenomena within the variable context. Poplack and Malvar (forthcoming), for instance, explain the favoring of Brazilian Portuguese PF in contexts without temporal adverbial modification as an indication of its default status as a future. While I agree with this assessment, I believe that the approach outlined here, which involves quantitative examination of some phenomena outside the variable context, will provide us with a wider range of possible interpretations for such differences.

7.4 Recommendations for further research

As mentioned in the Introduction, the SF and PF are not the only possible ways to express futurity in Spanish. Other forms are also in the mix—most notably the Present. Gutiérrez (1995), for example, does include this form in his study of the future tense in the Spanish of the Southwest United States. Likewise, Poplack and Malvar (forthcoming), examine the variation of four competing forms in Brazilian Portuguese, though with fewer tokens than were included here. An undertaking of such magnitude, which would include examination of all Present forms in the corpus, was not practical at this time. Nonetheless, further quantitative work is needed that takes into account both other forms that express futurity and other ways of expressing epistemic modality, in
order to achieve a fuller picture of the contexts in which these forms are found in variation.

Another item of interest that would be wise to take up is that of fossilization and fixedness in the SF (cf. Aaron 2003). As noted in Section 2.2, fixed phrases, as well as discourse-marker uses of ver, were excluded from all analyses. The reason for this exclusion was a posited uniqueness in distributional patterns, based on evidence from other fixed and semi-fixed expressions in spoken Spanish (Travis 2005, forthcoming). A quantitative study of the distributional patterns of these expressions (e.g. Duran and Gradoville 2006), especially from a diachronic perspective, could shed light on the connections between frequency, lexical diffusion, and discourse-marker formation.

A third site still in need of exploration is the development and competition of "backshifted" futures in iba a cantar 'was going to sing' and Conditional cantaria 'would sing'. This competition represents the past-tense version of the phenomena studied here, and was not considered. It would be most fascinating to compare these two semantically parallel developments.

7.5 Choosing a future

This study has built on the many decades of literature regarding the complex and ever-changing face of the Spanish future, and contributes more broadly to our understanding of the dynamics of variation during the process of grammaticization. Regarding the Spanish future, it is shown that the choice between the Periphrastic Future and the Synthetic Future has never, since the birth of the PF, been a question of one particular distinguishing semantic difference. Instead, this choice has been mediated by various factors, many of which have been suggested in previous literature. I was able to pinpoint
the changing distinguishing factors in PF-SF variation using overt contextual clues that were meant to operationalize notions such as intention and certainty.

Like other recent diachronic variationist work within a grammaticization framework (e.g. Torres Cacoulllos 2001; Poplack & Malvar forthcoming), I have provided quantitative empirical evidence for the process of grammaticization, most especially of generalization, in a localized, real-time context. Furthermore, the diachronic data provide clear support for Bybee, Pagliuca and Perkins' (1991) model for diachronic development of future forms. This evidence further shows that new forms do not spread across all contexts at the same rate, which is evidence against Kroch's (1989) Constant Rate Hypothesis, at least in cases of renewal; instead, this "spreading" is mediated by the lexical origins of the new form and the form(s) it is replacing (cf. Poplack & Malvar forthcoming for the same conclusion based on data from Brazilian Portuguese).

One of the fundamental questions of this and other similar studies has been simply: how do speakers choose a future form? We now have a handful of studies that have aimed at answering that question using variationist methodology. In Table 7.63, I have summarized the findings of the studies by Almeida and Díaz (1998) on Canary Island Spanish, Poplack and Turpin (1999) on Canadian French, Walker and Torres Cacoulllos (2005) on Canadian English, Poplack and Malvar (forthcoming) on Brazilian Portuguese, and the present study, on Peninsular Spanish, in order to discover if we have, indeed, come across any cross-linguistic tendencies in variation patterns within the realm of future temporal expression? This table presents the contextual features found to have a significant effect on the selection of (French, Portuguese, Spanish) SF, or in the case of English, will, the conservative form. In the case of Poplack and Malvar (forthcoming), I
have included the results for 19th century and 20th-century popular plays, since 20th-
century speech no longer includes the SF among the variants. Regarding my study, I have
included only the results for 20th-century speech.

Table 7.63. *Factors favoring the choice of SF vs. ‘go’-based future* in Romance
languages and English *will* (based on Almeida & Díaz 1998; Poplack & Turpin 1999;
Walker & Torres Cacoullos 2005; Poplack & Malvar forthcoming; and the present study,
Chapter 5)

|-----------|------------|------------|-------------|---------|----------|
| Stative/non-
motion verb | F         | F         | F           | [n.s.]  |          |
| Main clauses | F         | F         | [n.s.]      | F       |          |
| Non-specific temporal adverbial | F         | F         | F           | F       |          |
| Negative | D          |           | F           | D       |          |
| Polarity |            |           |             |         |          |
| Interrogative | D         |           | F           | D       | no consistent effect |
| Temporal distance: distal |            |           |             |         |          |
| Is subject | D          |           | D           | F       |          |

F = favors SF or *will*
D = disfavors SF or *will*
[n.s.] = factor not selected as significant
blank = no information provided

* Braz. Port. includes other variants besides PF.
** Other factors included only in Canary Island were unexpressed subjects (favor) and in Canadian French contingent (favor), imminent (not significant), and formal you (favor).

As we see in Table 7.63, three features show consistent favoring of the older form: stative
or non-motion verbs, main clauses, and nonspecific temporal adverbials. In the latter two,
English *will* lines up with Romance SF, while lack of verb class effect shows English *will*
lining up against Romance SF as a group. Given these similarities, it is tempting to
conclude that the clause type and adverbial effects that have been repeatedly found are
likely linked to *go*-futures, which are favored with subordinate clauses and specific
temporal adverbials or lack of adverbial modification. This may, indeed, be the case for
clause type, though no such conclusion can be made at this point. The temporal adverbial

265
effect, however, is not linked to go-futures in general. As seen in Section 6.3.3, the
temporal adverbial constraint appears to be linked to the (increased) association of SF
with nonspecific temporal adverbials. Similarly, the disfavoring of SF in interrogative
contexts is due to the decreased association of SF with intention, on the one hand, and
rhetorical incredulity, on the other (see Section 6.3.2). Walker and Torres Cacoullos also
find that this constraint is not linked to the go-future, but rather to will in English, due to
retention of volition meaning (2005). Further studies that operationalize measures in
similar ways will allow us improved accountability when comparing constructions cross-
linguistically. Nevertheless, the apparent consistency across varieties, languages, and
even language families (Romance and Germanic) is encouraging evidence that we are
moving toward a better understanding of universals through close examination of local
processes.

Through a combination of form- and function-based analyses of Spanish SF and
PF, I have proposed a methodology that aims to explain changes in constraints on
variation as reflections of shifts within the larger lexico-semantic space in which forms in
variation coexist, even when some of this space has not been shared by competing forms.
This proposed methodology includes analyses of results focused on both innovative and
conservative forms in terms of their respective paths of grammaticization (as in Torres
Cacoullos 2001—though forms aren't characterized as innovative and conservative in that
study), which by necessity includes form-based studies. The second step includes
quantitative analysis of those occurrences of the same forms that do not fall within the
variable context, however it has been defined, which may prove valuable in explaining
shifts in conditioning on variation.
This study has brought into the light a complex system, predicted in grammaticization theory, in which everything is delicately balanced, connected and in motion. Grammaticizing forms follow their natural paths toward semantic generalization and functional usefulness, all the while negotiating the division of labor with other forms that are capable of performing the same functions. And while forms may vie for default status within a contested functional space, this space is not immune to influence from nearby relatives, such as polysemies of competing forms, evidence of the depth and intricacy of what we call grammar.
Appendix: Corpus

Old Spanish

Mid-13th to mid-14th century


Late 15th century


Early 17th century


268
Late 18th to early 19th century


Late 20th to early 21st century, written corpus


Late 20th century, oral corpus

COREC = Corpus de Referencia de la Lengua Española Contemporánea: Corpus Oral

Peninsular, director F. Marcos Marín. Available at: www.illf.uam.es/~fmarcos/informes/corpus/corpusix.html (género conversacional).
References


parlé à Montréal, ed. Monique Lemieux and Henrietta Cedergren, 189-204. Québec: Gouvernement du Québec.


Jeanjean, Colette. 1988. Le futur simple et le futur périphrastique en français parlé. In *Grammaire et histoire de la grammaire: Hommage à la mémoire de Jean*


Moreno de Alba, J. 1978. *Valores de las formas verbales en el español de México.* Mexico City: UNAM.


Poplack, Shana and D. Turpin. 1999. Does the FUTUR have a future in (Canadian) French? Probus 11.133-64.

Powell, Barry B. Forthcoming. Writing, Blackwell.


Sankoff, Gillian. 1990. The grammaticalization of tense and aspect in Tok Pisin and Sranan. Language Variation and Change 2:3.295-312


Travis, Catherine E. Forthcoming. Subjetivización de construcciones: los verbos 'cognitivos' en el español conversacional. Serie Memorias del VIII Encuentro Internacional de Lingüística en el Noroeste.


