



Université d'Ottawa • University of Ottawa



# Université d'Ottawa - University of Ottawa

FACULTÉ DE ÉTUDES SUPÉRIEURES  
ET POSTDOCTORALES

FACULTY OF GRADUATE AND  
POSTDOCTORAL STUDIES

Elisabete Da Silva MALVAR

AUTEUR DE LA THÈSE - AUTHOR OF THESIS

Ph.D. (Linguistics)

GRADE - DEGREE

Department of Linguistics

FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

Future Temporal Reference in Brazilian Portuguese: Past and Present

S. Poplack

DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

CO-DIRECTEUR DE LA THÈSE - THESIS CO-SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

J. Jensen

M. Kato

A. Lapierre

E. Schneiderman

L.-M. De Koninck, Ph.D.

LE DOYEN DE LA FACULTÉ DES ÉTUDES  
SUPÉRIEURES ET POSTDOCTORALES

SIGNATURE

DEAN OF THE FACULTY OF GRADUATE  
AND POSTDOCTORAL STUDIES

# **Future Temporal Reference in Brazilian Portuguese: Past and Present**

**Elisabete da Silva Malvar**

Thesis submitted to the  
Faculty of Graduate and Postdoctoral Studies  
in partial fulfillment of the requirements  
for the Ph.D degree in Linguistics

Department of Linguistics  
Faculty of Arts  
University of Ottawa



National Library  
of Canada

Bibliothèque nationale  
du Canada

Acquisitions and  
Bibliographic Services

Acquisitons et  
services bibliographiques

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file* *Votre référence*  
*ISBN: 0-612-90004-5*  
*Our file* *Notre référence*  
*ISBN: 0-612-90004-5*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this dissertation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de ce manuscrit.

While these forms may be included in the document page count, their removal does not represent any loss of content from the dissertation.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

**Canada**

To the memory of my sister,  
Virgínia Malvar.

*Saudade*

## Abstract

---

Different forms convey future temporal reference in Portuguese. The recognition of these forms is a starting point towards examining their actual usage in the language. Of four acknowledged future forms, only two, i.e. the periphrastic *ir* + infinitive and the present variants, are productively used in contemporary spoken Brazilian Portuguese. There was therefore a need for reconstructing the future temporal reference system of early Portuguese, in order to shed light on the development of these forms in the future sector.

This study employs the methods of comparative and variationist linguistics in diachronic and synchronic data to analyse this development. These methods made it possible to identify the conditioning factors contributing to the selection of the variants in future contexts throughout the centuries and establish the direction of the linguistic evolution.

The first analyses concern the use of the future variants in the sixteenth and eighteenth centuries. They reveal that three variants were productively used to express the future: the synthetic future, *haver* + infinitive and the present. The synthetic future, which is acknowledged as the default form by prescriptive literature, had the highest overall probability of occurrence at that time, and was mainly favoured by contingent events, stative verbs and affirmative statements.

The second analyses concern the use of the variants in the nineteenth century: the synthetic future, *haver* + infinitive, the present and *ir* + infinitive. In that century, the synthetic future lost two of its favoured contexts, contingent events and affirmative utterances, to the present variant and *ir* + infinitive, respectively. That century witnessed the emergence of an adverbial specification pattern in Brazilian Portuguese that has also been observed in French.

The last analyses are related to the use of the future variants in the twentieth century. The results for the written data reveal that (i) *haver* + infinitive moved away from the future sector; (ii) the synthetic future declined; and (iii) the use of *ir* + infinitive highly increased. In the spoken data, only contingency and adverbial specification retained the present variant in the language. *Ir* + infinitive has almost completed its process of grammaticalization.

This research reveals that the future variants have been in variation for centuries in Portuguese, and have been involved in a complex process of incursion and withdrawal from the future context, which ultimately changed the representation of future temporal reference in contemporary Brazilian Portuguese. The variationist methodology makes it possible to ascertain that the difference in frequency of variant usage in the different centuries is correlated to their displacement of or expansion into specific linguistic contexts.

## Acknowledgments

---

I would like to take this opportunity to express my gratitude to those who encouraged me in my academic career.

I am deeply indebted to my supervisor, *Shana Poplack*, for her support all these years. The lessons I learned from her, specially her dedication to research, will always inspire me. I also extend my thanks to the members of my committee for accepting to review and comment this thesis.

My gratitude is extended to Professor *Eta Schneiderman* and those who taught me at the University of Ottawa, Professors *Denis Bouchard*, *Paul Hirschbühler*, *John Jensen* and *Maria Luisa Rivero*. I also would like to thank those who introduced me to the Variationist Theory, Professors *Stella Maris Bortoni-Ricardo* and *Maria Marta Scherre*.

My appreciation to *Yolande Thériault*, *Lise Picard*, *Jennifer Hill* and *Jeanne D'Arc Turpin* for their assistance in my academic life.

I acknowledge CNPq, University of Ottawa and CLDF for their financial assistance. In particular, I would like to express my appreciation to the Faculty of Graduate and Postdoctoral Studies (UO) as well as to *Ernesto de Paula* and *Jessy Pinheiro* (CNPq) for all their support.

This dissertation in particular has benefited from the support and infrastructure of the Sociolinguistic Laboratory, which provided me with workspace and opportunity to exchange ideas with a dynamic team of variationist sociolinguists.

I was blessed with the presence of two Brazilian Ph.D. students in Ottawa: *Vera Roloff* (Department of Linguistics) and *Rossana Andrade* (Department of Computer Science). They gave me emotional support, made me laugh and enjoy life far from home. Thanks also go to my colleagues, particularly, *Olga Arnadouva*, *Rodica Diaconescu*, *Nathalie Dion*, *Ejike Eze*, *Dawn Harvie*, *Carmem LeBlanc*, *Martine Leroux*, *Ahmad Moinzadeh*,

*Reza Ghafar Samar, Rebecca Silvert, Danielle Turpin, Gerard Van Herk and James Walker*, as well as to my friends in Ottawa, specially, *Káthia and Nicholas Anquetil, Adriana and Don Corcoran, Pérpetua and Pierre Landry, Lígia and Riverson Rios, Viviane and Frank, Sandra Clark*, and *Cristiane and Gabriel Misaka* for the good time we spent together.

I am indebted to *Rena Torres Cacoullós* for sharing her office with me and for bringing life to that hard winter.

*Rebecca Silvert* has far exceeded what is expected from a colleague. I appreciate her careful reading and valuable comments on the last draft of this dissertation. The usual disclaimers apply, of course; I am responsible for remaining imprecision or mistakes.

I am fortunate in my family and friends. I would like to express my eternal gratitude to my parents, *Raimundo and Geni Malvar*, to my sister, *Regina*, to my nieces *Bárbara, Bruna and Brenda*, and to my husband, *Eliézio da Silva*. Their patience, encouragement and love gave me enough strength to stay abroad and to work on my research.

A warm thank you goes to my cousins *Jorge, Jurema, Helena, Henrique*, and *their families*, to my godmother *Gilséa Malvar* and to my uncle *Lincoln Silva*. Sadly, my dear aunt *Helena Malvar* and my godfather *Henrique Malvar* passed away in the course of my program. They are always with me.

Last but not least, my special gratitude goes to my informants and to my friends in Brazil, specially, *Marcony Vinícius, Virginia Fátima, Rosana, Iara, Cecília, Nilson, Zita, Cléa, Nóe, Gláucia and their families* for their prayers, mails, calls, and encouragement.

It has been hard to get through these years without my sister *Virginia Malvar*. This dissertation is dedicated to her.

# Table of Contents

---

Abstract.....	iii
Acknowledgments.....	v
Table of Contents .....	vii
List of Figures .....	x
List of Tables.....	xi
<b>CHAPTER 1 INTRODUCTION.....</b>	<b>01</b>
1.1. Future Temporal Reference in Portuguese .....	01
1.2. The Present Study .....	04
1.2.1. Method.....	06
1.2.2. Data.....	08
1.3. Justification for the Present Study .....	09
1.4. Overview of the Dissertation .....	09
<b>CHAPTER 2 THE FUTURE.....</b>	<b>11</b>
2.1. Some Words on the Future in Latin.....	11
2.1.1. The Meanings of the <i>Habere</i> Periphrasis.....	14
2.2. Future Forms in Early Portuguese .....	16
2.3. Tracing the Meanings of the Future Forms in Portuguese Grammar Books .....	18
2.3.1. Future Forms in Grammar Books of the Sixteenth to the Eighteenth Centuries.....	26
2.3.1.1. Mentions of Variants.....	26
2.3.1.2. Mentions of Meanings.....	26
2.3.2. Future Forms in Grammar Books of the Nineteenth Century.....	30
2.3.2.1. Mentions of Variants.....	30
2.3.2.2. Mentions of Meanings.....	30
2.3.3. Future Forms in Grammar Books of the Twentieth Century.....	32
2.3.3.1. The Synthetic Future .....	33
2.3.3.2. <i>Haver</i> + Infinitive.....	34
2.3.3.3. The Present Form .....	37
2.3.3.4. <i>Ir</i> + Infinitive.....	40
2.4. Summary .....	41
2.5. Future Form Meanings in Prescriptive Works and their Handling by Variationist Research.....	43
<b>CHAPTER 3 THEORETICAL FRAMEWORK, METHOD AND DATA.....</b>	<b>45</b>
3.1. Theoretical Framework and Methodology.....	45
3.1.1. The Form – Function Asymmetry.....	48
3.2. The Scope of the Investigation .....	50
3.3. The Data Set .....	51
3.3.1. The Diachronic Data.....	51
3.3.1.1. The Selection of Plays.....	54
3.3.1.1.1. Portuguese Plays.....	54
3.3.1.1.2. Brazilian Plays .....	58
3.3.2. The Synchronic Data .....	65
3.3.2.1. Brasilia .....	65
3.3.2.2. The Interviews.....	67
3.3.2.3. The Sample .....	68

<b>CHAPTER 4 ANALYSIS .....</b>	<b>71</b>
4.1. The Linguistic Variable .....	71
4.2. Circumscribing the Variable Context .....	74
4.3. Previous Quantitative Works on the Future Forms in Brazilian Portuguese .....	77
4.4. Linguistic Hypotheses, Factor Groups and Coding .....	80
4.5. Some Considerations on the Groups of Factors.....	92
<b>CHAPTER 5 RESULTS .....</b>	<b>97</b>
5.1. Overall Distribution of the Variants .....	97
5.2. Distribution of Variants by Century .....	98
5.3. The Sixteenth and Eighteenth Centuries.....	99
5.3.1. The Sixteenth and Eighteenth Century Variable Rule Analyses.....	102
5.3.1.1. Factors.....	102
5.3.1.2. Results.....	103
5.3.1.2.1. The Synthetic Variant .....	105
5.3.1.2.2. The Periphrastic <i>Haver</i> + infinitive Variant .....	107
5.3.1.2.3. The Present Variant .....	109
5.4. The Nineteenth Century.....	112
5.4.1 The Synthetic Variant .....	115
5.4.2. The Periphrastic <i>Haver</i> + infinitive Variant.....	117
5.4.3. The Present Variant .....	119
5.4.4. The Periphrastic <i>Ir</i> + infinitive Variant.....	121
5.5. Comparison between the Sixteenth and Eighteenth as well as the Nineteenth Century .....	124
<b>CHAPTER 6 RESULTS FOR THE TWENTIETH CENTURY.....</b>	<b>131</b>
6.1. Variable Rule Analyses of the Twentieth Century Data.....	132
6.1.1. Factors.....	133
6.1.2. Results.....	133
6.2. Written Data Analyses.....	137
6.2.1. The Synthetic Future.....	137
6.2.2. The Present .....	140
6.2.3. The Periphrastic Variant <i>Ir</i> + infinitive.....	146
6.3. Comparison of the Results of the Written Data: 16 <sup>th</sup> – 19 <sup>th</sup> Centuries vs. 20 <sup>th</sup> Century.....	149
6.4. Variable Rule Analyses of the Spoken Data.....	152
6.5. Comparison: Twentieth Century Written and Spoken Data .....	157
6.6. The Trajectory of the Periphrastic Variant <i>Ir</i> + infinitive in the Variable Context.....	162
6.7. Prescription vs. Usage .....	169
6.7.1. Meanings attached to the future variants .....	170
6.7.2. Comparison of the Distribution of the Variants in Formal/Informal and Writing/Speech.....	172
6.7.3. Level of Education.....	173
<b>CHAPTER 7 CONCLUSION .....</b>	<b>178</b>
7.1. Findings of this Study.....	180
7.1.1. The Future Variants .....	180
7.1.1.1. The Synthetic Future .....	181
7.1.1.2. <i>Haver</i> + infinitive .....	184
7.1.1.3. The Present Variant .....	185
7.1.1.4. <i>Ir</i> + infinitive .....	187
7.1.2. The Value of Plays .....	189
7.1.3. Prescriptivism .....	191
7.2. Conclusion .....	192

<b>REFERENCE .....</b>	<b>194</b>
<b>APPENDIX 1 - GRAMMARS, USAGE MANUAL, PRESCRIPTIVE LETTERS .....</b>	<b>204</b>
<b>APPENDIX 2 - PLAY REFERENCES .....</b>	<b>207</b>

# List of Figures

---

Figure 2.1: Percentage of Times Meanings are Ascribed to Future Forms ..... 25  
Figure 6.1: Speakers' level of education..... 174  
Figure 7.1: Usage of the future variants throughout the centuries in Portuguese..... 179

## List of Tables

---

Table 2.1: Consulted works.....	20
Table 2.2: Summary of the meanings attached to the future forms in usage manual, prescriptive and descriptive grammars of Portuguese by year (16 <sup>th</sup> to 20 <sup>th</sup> century)	22
Table 3.1: Selected Plays .....	64
Table 3.2: Informants.....	70
Table 4.1: Distribution of future variants in formal spoken Brazilian Portuguese. (adapted from Baleeiro 1988).....	77
Table 4.2: Distribution of future variants in formal written Brazilian Portuguese. (adapted from Santos 1997: 45).....	78
Table 4.3: Cross-tabulation of Grammatical Person and Animacy Group of Factors.....	92
Table 4.4: Distribution of the main verb <i>ir</i> 'to go' in future temporal reference contexts in the Portuguese corpus by variants. ....	94
Table 4.5: Occurrence and distribution of verbs with the present variant in the corpus. (N = number of occurrences).....	95
Table 5.1: Overall distribution of the variants of the future temporal reference in Portuguese.....	97
Table 5.2: Distribution of the variants of Future Temporal Reference in Portuguese by century (all tokens) .....	98
Table 5.3: Distribution of the variants of Future Temporal Reference in Portuguese after combining the Sixteenth and the Eighteenth century data.....	100
Table 5.4: Three variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, <i>Haver</i> + infinitive and Present will appear in future temporal reference contexts in the sixteenth and eighteenth centuries in Portuguese. ....	104
Table 5.5: Four variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, <i>Haver</i> + infinitive, Present and <i>Ir</i> + infinitive will appear in future temporal reference contexts in the nineteenth century in Brazilian Portuguese. Without the main verb <i>ir</i> 'to go'.....	114
Table 5.6: Seven independent variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, the periphrastic <i>Haver</i> + infinitive, the Present and the periphrastic <i>Ir</i> + infinitive will be selected in future temporal contexts in Brazilian Portuguese.....	127
Table 6.1: Distribution of the variants of the Future Temporal Reference in Brazilian Portuguese by written and spoken data from the twentieth century .....	133
Table 6.2: Distribution of the variants of the Future Temporal Reference in Brazilian Portuguese according to stativity in the twentieth century. ....	134
Table 6.3: Three variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, the Present and the periphrastic variant <i>Ir</i> + infinitive will appear in future temporal reference contexts in the twentieth century in Brazilian Portuguese. Without the main verb <i>ir</i> 'to go'.	136

Table 6.4: Three variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, the Present and the periphrastic variant <i>Ir</i> + infinitive will appear in future temporal reference contexts in the twentieth century in the written data. Without the main verb <i>ir</i> 'to go' .....	138
Table 6.5: Relative weights of the variants in main and subordinate factors in the written data from the sixteenth to the twentieth centuries. ....	143
Table 6.6: Ten independent variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, the periphrastic variant <i>Haver</i> + infinitive, the Present and the periphrastic variant <i>Ir</i> + infinitive will be selected in the future temporal context in Portuguese. Brackets indicate that the factor group was not selected as significant.....	150
Table 6.7: Variable rule analysis of the contribution of factors selected as significant to the probability that the Present variant will appear in future temporal reference contexts in contemporary spoken Brazilian Portuguese. Without the main verb <i>ir</i> 'to go' .....	154
Table 6.8: The trajectory of the Present and the periphrastic variant <i>Ir</i> + infinitive with the factor groups selected as statistically significant in contemporary spoken Brazilian Portuguese.....	155
Table 6.9: Five independent variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, the Present and the periphrastic <i>Ir</i> + infinitive variant will be selected in future temporal context in Brazilian Portuguese in the twentieth century. (Brackets indicate that the factor group was not selected as significant).....	159
Table 6.10: Distribution of future variants according to three studies with different corpora in twentieth century Brazilian Portuguese. ....	172

# Chapter 1 Introduction

---

“*Vou escrever* (GP)<sup>1</sup> um livro que se chama *A Arte da Política*, título provisório. *Vou pegar* (GP) alguns clássicos e *vou juntar* a minha experiência no governo para escrevê-lo. Mas o ano de 2003, *vou começar* (GP) na Europa. Logo após a posse, *embarco* (P) para a França, depois *vou* (P) para a Espanha, Itália, Inglaterra e Estados Unidos, países onde tenho alguns compromissos. Espero fazer tudo isso em seis meses. No segundo semestre *volto* (P) para São Paulo e *sento* (P) para escrever o livro”<sup>2</sup>

‘*I’m going to write* a book called *A Arte da Política*, as a provisory title. *I’m going to refer* to the Classics and then add my government experience to the writing of it. Then in 2003, *I will begin* my life in Europe. Immediately after the ceremony, *I will go* to France, then Spain, Italy, England and the United States, all countries in which I have some obligations. I hope to do all of this in six months. In the second semester, *I will come* back to São Paulo, and *will sit* and write the book.’

## 1.1. Future Temporal Reference in Portuguese

The future in Portuguese is expressed by four main forms, exemplified in (1) to (4)

below:

### (1) Synthetic Future

Esopo, desavergonhado, tu me foste enganar? Pois em ti *vingarei* (SF) minha raiva.

(Silva 1734: VE: Euripedes: second part, scene 3: 106)<sup>3</sup>

‘Esopo, shameless one! Did you deceive me? I *will get back* at you.’

---

<sup>1</sup> The symbols in parenthesis refer to the variants as follows: SF stands for *synthetic future*, HP stands for the periphrastic *haver + infinitive*, P stands for the *present* variant and GP stands for the periphrastic *ir + infinitive*. All forms used in future temporal contexts.

<sup>2</sup> Quoted from an interview with the former President of Brazil, Fernando Henrique Cardoso, in a well-known magazine in Brazil (*Isto é – Gente*), December, 2002.

<sup>3</sup> The codes in parenthesis refer to the author, the year of the play, the code of the play, character, location and page number. All examples taken from the oral data, as in (3), will have codes in parenthesis denoting the year of the interview, the informant, the side of the tape, as well as the count or the page number.

(2) *Haver* Periphrasis

Descanse, Senhor, que D. Nise *há-de ser* (HP) sua, apesar das cautelas do tio e das carícias do noivo. (Silva 1737: AM: Fagundes: scene 1: 48)

‘Relax, Sir, because D. Nise *will be* yours, despite her uncle’s protection and her fiancé’s affections.’

## (3) Present

A gente tem nosso grupo. É às oito da noite, na sexta. Amanhã *tem* (P) reunião. (1992: Vânia. B: 38)

‘We have our group, at eight o’clock, on Fridays. There is *going to be* a meeting tomorrow.’

(4) *Ir* Periphrasis

Os acessórios estão velhos? Pode ser, *vou pensar* (GP) no caso. Se as reivindicações são justas, serão atendidas. Taí, gostei. *Vou organizar* (GP) um sindicato pra vocês.

(Buarque 1978: OM: Duran: scene 3: 79)

‘The accessories are old? That may be. I *will think* about it. If the requests are fair, they will be considered. Oh, I know! I *will organize* a trade union for you.’

Of the four forms, prescriptive works consider the synthetic future, as in (1), to be the one that properly represents the future tense in Portuguese. Consequently, this is the form that appears on conjugation tables and that is propagated by the educational system. The other future forms assume peripheral status. They are generally considered to be used only when speakers wish to convey other meanings besides futurity. Yet, there is no consensus as to their precise meanings. Attention to spoken language will reveal that the

expressions of futurity exemplified in (1) and (2) are rarely, if ever, heard in future temporal contexts in Brazilian Portuguese. It is the future forms in (3) and (4) that are used in the spoken language.

Sociolinguistics discovered (e.g. Poplack and Turpin 1999, Poplack et al. 2002) that this contradictory situation of a form being posed as the standard one, yet not actually being used in speech, is not unusual. Nor is it unusual, when one linguistic function is represented by more than one form in a language, for prescriptive works to assume one form as the standard usage and to label the others as incorrect, or as expressing additional meanings, related to the speakers' intuition, feelings or interpretation of reality. This approach to language is rooted in the notion of form/function symmetry (i.e. to each linguistic form corresponds one function) common in contemporary linguistics. Consequently, variation is seen as synonymous with an inefficient language system and with obliteration of a desired stability. However, as stated by Labov (2001):

The first contribution of sociolinguistic research in the second half of the 20<sup>th</sup> century was to show that variation was not chaotic, but well-formed and rule-governed, that it was indeed an aspect of linguistic structure. (Labov 2001: 38)

This view of language as rule-governed and inherently variable provides an answer for what is observed in the future temporal reference sector in Brazilian Portuguese. The situation is seen as one of inherent variability involving all of the forms exemplified in (1) – (4). The question that arises is how exactly the future forms are partitioning the future temporal reference in the language and which rules are determining their choice in speech.

More specifically, how have the future forms that were exemplified in (1) to (4) been selected in the future temporal reference context throughout the centuries, given that (3) and (4) are now the future forms competing for this context in spoken Brazilian Portuguese.

## 1.2. The Present Study

This study is designed to answer the question: how have the future forms been competing for the future temporal reference in Brazilian Portuguese? To determine precisely how these future forms were introduced in the Portuguese language and how they have been used entails diachronic investigation. This investigation is twofold: (i) investigation of the meanings that were ascribed to the future forms in historical and prescriptive grammar books, and (ii) investigation of the actual usage of the future forms in the language.

First, an inventory of grammar books was inspected, starting with those of the sixteenth century, that being when the first grammar books were published in Portuguese. Every remark on the usage of the future forms was recorded, along with its associated meanings. As shall be revealed in chapter 2, prescriptive literature has shown an inability to deal with the dynamic process of usage of future forms in Portuguese throughout the centuries. The remarks of these grammarians are mainly concerned with prescribing the synthetic future as representing futurity in the language. Also, they propose a list of meanings to the other future forms, which are related to the speakers' engagement or view of future eventuality.

The goal of this investigation is (i) to find out potential factors that would account for the selection of each future form in the variable context; then, (ii) to relate these factors to specific hypotheses; and (iii) to test these factors by means of variable rule analyses.

Every prescriptive remark regarding the meanings or usages of the future forms that could be operationalized as a factor was incorporated in the quantitative analyses.

The diachronic investigation of the usage of the future variants required data in Portuguese from previous centuries. Since the Variationist Theory relies on data from speech, there existed the hard task of finding trustworthy sources of data from periods of time when there were no records of spoken language. Chapter 3 details the gathering of the data. Careful research was done into literary genres, individual authors' styles and linguistic features of their works. Popular plays were selected as the literary works most representative of Portuguese speech.

Variable rule analyses of texts from the sixteenth and eighteenth centuries<sup>4</sup> and the nineteenth century show that although all the variant forms are involved in the future temporal context, they are involved in a dynamic process of incursion into and withdrawal from this context. Their selection is not consistent with all of the meanings attributed to them by Portuguese prescriptive literature. Variable rule analyses of these first two centuries studied show that primarily the synthetic future and the periphrastic form *haver* + infinitive competed for the variable context at that time, being mainly constrained by the factor of *contingency*. The next century witnessed greater competition for the future temporal reference sector. All four future variants are engaged in this competition, being mainly constrained by *adverbial specification*, *contingency*, *stativity* and *polarity*. Chapter 5 depicts the results of the variable rule analyses of texts from these centuries.

Subsequent to these analyses, a synchronic analysis was designed to investigate the future temporal reference sector in twentieth century plays and in the data collected from Brazilian Portuguese speakers. This inspection made it possible to trace the usage of the

---

<sup>4</sup> Data from the seventeenth century are not available. Please refer to chapter 3 for details.

variants from previous centuries to the twentieth century. Moreover, it was also possible to track their usage throughout the twentieth century, and to compare it with that of both the examined plays and the speech of the end of that century. It was found that in twentieth-century Brazilian Portuguese, almost all reference to future eventuality is made by the periphrastic form *ir* + infinitive. The results show that by the end of this century the usage of this variable is constrained only by *contingency* and *adverbial specification*, which are the linguistic factors that favour the present form usage in the future context.

This research on spoken Brazilian Portuguese was also designed to examine whether the prescriptive tradition, mainly perpetuated by the educational system, affects the usage of these variants in the language. Hence, the informants were classified by their different levels of education. One would have thought that the higher their level of formal education, the more they would have been likely to select the synthetic future in the future context. However, the results show that, as far as the usage of future variants is concerned, university-level informants follow a similar pattern to that of the elementary-level informants. Thus, I conclude that the prescriptive tradition has not affected the course of natural language.

### *1.2.1. Method*

As mentioned in the previous section, this dissertation adopts the Variationist Methodology as originally established by Labov (1966, 1969, 1972a). This methodology assumes that variable phenomena must be analyzed in linguistic data which present regularity in the production of linguistic contexts. Isolated examples or intuitive inferences on the usage of linguistic variants do not provide an accurate view of either the processing

involved in the selection of variants by speakers, or the constraints that favour this selection. Investigating variation presupposes (i) meticulously circumscribing the variable context, i.e. the context in which all forms representative of a phenomenon are equally allowed without change in meaning; (ii) extracting all instances of variable context from the data; (iii) postulating hypotheses, which may involve linguistic or extra-linguistic factors; (iv) coding each instance of variation according to these factors; and (v) performing multivariate analyses in order to ascertain which factors contribute the most to the selection of variants. This method enables researchers to determine how variants are chosen as well as how they are entrenched in variable contexts.

This research also makes use of the Comparative Method as established in Poplack and Tagliamonte (2001: 89-100). Based on this method, which accommodates the notion of inherent variability, multivariate analyses of different stages of language may undergo comparison. After verifying the effect of the conditioning factors on the variants in different periods of language, the constraint hierarchies that emerge from the variable rule analyses may be compared. In the present study, this comparison is essential for identifying the extent to which linguistic factors have consistently affected selection of the future variants throughout the centuries, and in particular, the way in which the periphrastic variant *ir* + infinitive takes over the linguistic contexts that previously favoured the occurrence of other variants.

### *1.2.2. Data*

This study is grounded in empirical investigation, namely every conclusion is supported by data; therefore, the gathering and the treatment of data is regarded as central to the research. As established in variationist methodology, the most reliable data for analysis are obtained through tape-recorded interviews that capture native speakers' everyday usage of language. To this end, I planned and carried out interviews with native speakers of Brazilian Portuguese. The interviews followed the sociolinguistic model in order to overcome the limitations of tape-recorded interviews. Using this dataset, variation in the future temporal reference sector in Brazilian Portuguese at the end of the twentieth century was analyzed.

Preliminary observations of natural speech and grammatical prescriptions led me to the question of how different the usage of the future variants was in the past; i.e., whether this striking contradiction between speech and prescription is a result of change in the future sector in the Portuguese language. Accordingly, this study incorporates data from the Portuguese written language. The data were gathered from popular plays, which proved to be representative of Portuguese speech. The data set consists of twenty-two plays, written over four different centuries. The written data made it possible to study the variation in the future temporal sector in early Portuguese.

Chapter 3 gives further details on the methodology and data set used in this study.

### 1.3. Justification for the Present Study

The variationist perspective was used in a recent investigation of the future temporal reference sector in the French and the English languages (Poplack & Turpin 1999, Poplack & Tagliamonte 2000, Poplack & Tagliamonte 2001). In Brazilian Portuguese, this variation was examined by: (i) Baleeiro (1988) in the formal spoken language of São Paulo, using a quantitative approach, but without performing variable rule analyses; and (ii) Santos (1997) in the formal written language of parliament discourses, following the standard variationist approach. Both these studies are summarized in chapter 4.

Although the above studies of the Brazilian Portuguese language examined the same variable that is the focus of this dissertation, no study has emerged regarding the issues that inspired me to pursue this research: (i) the variation in spoken colloquial language; (ii) the variation in the early stages of the language; (iii) the incongruity between prescription and usage; (iv) the comparison of the effects of the conditioning factors throughout the centuries; and (v) the grammaticalization process of the periphrastic future variant *ir* + infinitive. An analysis of these issues will contribute to the understanding of this long-term variable phenomenon in Portuguese and any change that has occurred in this sector. In addition, it will broaden our understanding of the expression of Future Temporal Reference in Romance languages more generally, most of which have at their disposal at least some of the variants treated here.

### 1.4. Overview of the Dissertation

The outline of this work is as follows. Chapter 2 presents the search for the origin of the future forms in Portuguese. It is centered on the investigation of the treatment given to

these forms by prescriptive works throughout the centuries. It provides both a summary of the meanings associated with each future form, and an explanation of the way in which variationist research handles prescriptive remarks.

Chapter 3 establishes the methodology and the dataset used in this study. It focuses on the constitution of the data, that is, the parameters for data collection and the research that validated the plays as representative of speech. Chapter 4 is designed to (i) circumscribe the variable context in the language; (ii) elaborate and specify the hypotheses that will be tested in the variable rule analyses; and (iii) present the overall distribution of the data.

Chapters 5 and 6 discuss the variable rule analyses. It provides variable rule analyses of the future temporal reference variation in the sixteenth and eighteenth centuries as well as the nineteenth. It then presents the comparison of the results. Chapter 6 exhibits the variable rule analyses of the twentieth century. It focuses on two lines of inquiry: the results of the variable rule analyses for written representations of speech throughout the centuries, and a comparison of such representations with tape-recorded interviews in the twentieth century. This chapter examines the grammaticalization of *ir* + infinitive in the language and, in order to investigate the influence of the prescriptivism in speech, also discusses the degree of variation in the speakers' language according to their level of education.

Finally, chapter 7 summarizes this study's major findings as well as their importance to the understanding of the future temporal reference sector in Brazilian Portuguese.

## Chapter 2 *The Future*

---

The nature of the future as a mental construct has been the object of a great number of studies and of much speculation. The temporal location of events and actions expressed without a real and material existence are commonly regarded as a result of hope, desire, expectation, prospect, probability, obligation or intention. In order to understand how the future is mapped onto the grammatical category tense in Portuguese, a few words on the linguistic representation of the future in Latin and Portuguese are presented in this chapter. The main concerns here are (i) the meanings attributed to the forms that are conceived of as expressing future, and (ii) how some linguists, as well as historical, descriptive and prescriptive grammarians assert the usage of more than one form as exponents of future temporal reference in Portuguese.

### 2.1. Some Words on the Future in Latin

In Classical Latin the future temporal reference paradigm was represented by - *bo*, e.g. *amabo*, for the first and the second verbal conjugations, and by -*am*, for the third and fourth ones, e.g. *legam*<sup>5</sup>. These representations are said to have evolved into the Latin future from an archaic periphrastic future that had volitional meaning in the case of the -*bo* inflection, and from the subjunctive present in the case of the -*am* form. It is also thought that the -*bo* suffix was formerly the present form of the verb ‘to be’, which was added to the stem of the verbs of all conjugations, and only later was excluded from the third and

---

<sup>5</sup> Example of future inflection for 1<sup>st</sup> and 2<sup>nd</sup> regular conjugation: *amābo* ‘I will love’, *amābis*, *amābit*, *amabīmus*, *amabītis*, *amābunt*. Example of future inflection for 3<sup>rd</sup> and 4<sup>th</sup> conjugation: *legam* ‘I will read’, *leges*, *leget*, *legēmus*, *legētis*, *legent* (cf. Almeida, 1985: 208-222).

fourth conjugations by the *-am* form (Mattoso Câmara 1972: 110, 1980: 124; Fleischman 1982: 34).

This Classical Latin future form did not prevail in Vulgar Latin (Ferreiro 1996: 296; Mattoso Câmara 1967: 30; Nunes 1910: 270). There are certain crucial linguistic reasons for the ruling out of the *-bo*, *-am* future forms. The first is the sound change of the intervocalic /b/ into /v/, which caused an overlapping of forms between the future, and the first and second conjugation verbs in the past tense, e.g. *amābit* in the future and *amāvit* in the past. The second is the similarity of the third and fourth conjugation future forms with those of the subjunctive, e.g. *legam* and *capīam*. Moreover, similarities of the third and fourth conjugation forms with the indicative present are claimed to be yet another cause for their exclusion (Mattoso Câmara 1967: 30).

Although there were several future forms in Latin, they are commonly said to have occurred sequentially. However, we need to keep in mind that they did co-exist over a substantial period of time. For instance, periphrastic constructions were used to express the future in Classical and in Later Latin: the combination of the future active participle of a verb with synthetic forms of ‘to be’, e.g. *uenturus sum*, and also the future passive participle with these same forms, e.g. *amandus sum*, are cited as having expressed futurity regardless of an inventory of psychological distinctions ascribed to them. Thus, the *uenturus sum* type implies imminence, intentionality, destiny, and resolution; while the *amandus sum* type retains the meanings of intent or goal, in addition to possibility, obligation and necessity (cf. Mattoso 1967: 27; Fleischman 1982: 35-39). In later Latin, futurity was more generally ascribed to the gerundive construction.

Neither the synthetic future form, nor any of these periphrastic constructions were maintained in Romance. A periphrasis formed by the imperfect of the infinitive and the present of *habere* 'to have' is attested as having been used in Vulgar Latin, and as having been the one which was spread to the Romance-speaking areas (Nunes 1910: 270; Ferreiro 1996: 296; Mattoso Câmara 1972: 110; Said Ali 1921/ 2001: 111; Silva Neto 1957/ 1986: 254). Silva Neto (1957/ 1986: 255), for instance, cites examples of *habere* periphrasis in Tertullian (A.D. 160-230), the first theologian to register the construction, yet in these examples the periphrasis is seen to imply obligation.

Mattoso Câmara (1972) argues that the use of a future tense was not favoured in Vulgar Latin because a strictly temporal future concept did not operate in the colloquial usage of Indo-European languages. Thus, in spontaneous communication the notion of future is always associated with other meanings, such as doubt, desire or imposition of will. Later in the language, "the grammatical discipline and the norm of written (or literary) language were able to deflect these forms toward a strictly temporal future" (Mattoso Câmara 1972: 110). This process is also observed in Classical Latin in which the future tense arose from periphrastic forms. Mattoso Câmara claims that "on all levels of the social hierarchy of speakers of Vulgar Latin, the predominant tendency was the replacement of future tense by the present, unless there was some specific motivation to the contrary" (Mattoso Câmara 1972: 110).

Mattoso Câmara's statement favours the use of non-synthetic future forms in spoken Latin, and in a broader sense in all Indo-European languages. However, it can be inferred from his words that any form, used in a future temporal reference context, conveys meanings other than the temporal future if it has not been previously accepted in the written

language. Analogous reasoning is cited by Silva Neto (1957/ 1986: 255), although he gives an historical explanation. He assumes that the *habere* periphrasis arose in the language because a change in people's mentality occurred due to religious worries and to the decline of Classical Latin culture, both of which were translated into superstitious feelings among the popular classes. As a consequence, people lost the objective way of seeing their future and began to express their uncertainty, fear and hope by means of periphrastic constructions. However, since the synthetic future form in Classical Latin also resulted from a periphrastic construction, such a statement is unreasonable unless one assumes that all linguistic structures result from superstitious feelings among the popular classes.

In the search for the meanings attributed to the future forms, it was found that statements like these seem to be a generalized preconception maintained by grammarians, as well as some linguists. The more the future forms are associated with the spoken language, the more they are considered to be carrying special meanings which are mainly correlated to the psychological and emotional attitudes of speakers, i.e. meanings which cannot be observed. Moreover, these views are substantiated by a few examples, which are commonly analyzed in such a way as to convince speakers that they chose a particular future form because of a specific intention they had in mind, other than futurity. In the next section, specific meanings attached to the *habere* periphrasis in its emergence will be analysed.

### *2.1.1. The Meanings of the Habere Periphrasis*

The *habere* periphrasis was first attested in a written Christian document of the third century. However, there is no consensus in the literature on the meaning that it had at

that time. Analyses of sermons and documents suggest that it did not assume a future/posterior meaning until the eighth century. But, again, this is controversial since Taylor (1924, cited by Fleischman 1982: 53) claims that there is evidence of a co-existence of the *-bō* synthetic future and *habere* periphrasis (analytic *cantare habeo/ habeo cantare*) in the eighth century in Alcuin's letters, in which, he insists, the periphrasis expresses obligation and maintains a possessive meaning in many examples. Ferreiro (1996: 297), Maia (1986: 764) and Mattoso Câmara (1972: 148) also ascribe the meaning of obligation to this construction, hence, the periphrasis is called the *obligatory* or *necessitative* future<sup>6</sup>.

Mattoso Câmara (1967: 26; 1972: 111) also suggests that the *habere* periphrasis was a volitional expression which speakers could use to stress their desire regarding the occurrence of a certain event, and because this desire was anchored in a present perspective, the auxiliary was in the present tense. According to Câmara, the motivation for a mood-like future, which in Classical Latin promoted the volitional form *-bo* to future, was still at work in Vulgar Latin. By means of the *habere* periphrasis, speakers could declare, for instance, that they want to sing, intend to sing or have singing as a goal at the present moment.

Another interpretation for this periphrasis is found in Benveniste's article (1968). The author claims that the *habere* periphrasis displayed an idiosyncratic syntactic structure and by no means opposed the conventional future. It was used in subordinated structures and functioned as a verbal adjective or a participle, since it mainly appeared in relative clauses. Thus, this periphrasis was equivalent to a future passive participle. However,

---

<sup>6</sup> In Portuguese, obligatory/ necessitative meaning is conveyed by a construction of the following type: the auxiliary *ter* 'to have' + the preposition *de* 'of' + an infinitive, e.g. *tenho de cantar* 'I have to sing'. As in English, this construction is used in all tenses.

Benveniste asserts that it did not signify obligation, but rather denoted a sense of *destiny* or *inevitability*, namely, its meaning was *predestination, in verbis*:

In the Latin syntagm as it actually crystallized *habere* + infinitive served to indicate the predestination of the object to follow a certain course of events. This is a novel and distinctive semantic hue, totally divorced from the purposive value often associated with the future tense. (Benveniste 1968: 90)

According to Benveniste, the periphrasis first extended its hold to independent clauses, then *habere* started to be combined with intransitive verbs, and finally it became combined with the infinitive of all verbs. Only in the sixth and seventh centuries did its usage become generalized. From that time onwards, the periphrasis competed with the traditional future (*-bō, -am* forms), and eventually supplanted it. In sum, according to the author, the periphrasis signified a sense of fate, i.e. a total absence of will, rather than a sense of obligation, possessiveness or volition.

## 2.2. Future Forms in Early Portuguese<sup>7</sup>

The *habere* periphrasis followed the same path in almost all Romance languages<sup>8</sup>. The *habere* + infinitive/ infinitive + *habere* constructions co-existed and were incorporated into these languages. Then, *habere* suffered phonetic reduction and the periphrasis preceded by the infinitive underwent an agglutination (Silveira Bueno 1967: 142; Cardoso & Cunha 1978: 120; Huber 1933: 230; Mattoso Câmara 1972: 112; Nunes 1910: 319;

---

<sup>7</sup> Portugal established itself as an independent kingdom in 1143. The first documents attested in the language dated back to the end of the eleventh and the beginning of the twelfth centuries (Huber 1933: 23).

<sup>8</sup> Nunes (1910: 319) observes that the Latin verb *habere* was replaced by *velle* and *venire* in Valaquo and Ladino, respectively.

Williams 1968: 207). In Portuguese, *amar hei* ‘I will love’, for instance, became *amar ei* and later *amarei*; whereas *hei amar* circulated with the preposition *de* and became *hei de amar*. Nunes (1910: 320) claims that the placement of the preposition in the periphrasis was artificial at first because it was only inserted into the written language, whereas *hei amar* was mostly used and maintained in the language by popular speech. Concerning the new synthetic form, the precise date of its emergence in Portuguese is not mentioned. As is well-known, linguistic change is not immediately reflected in the written variety. Moreover, in early Portuguese the spoken and written languages were extremely different because the latter was mainly registered in judicial documents. Maia (1986: 841) however, attests the use of the synthetic future form of *ter* ‘to have’ as *teerei* and *terei* in the *Cancioneiro da Vaticana* and in the *Cancioneiro da Ajuda*, respectively. In addition, Said Ali (1921/ 2001: 111) cites the use of the synthetic future forms in the Portuguese manuscript entitled *Demanda do Santo Graal*<sup>9</sup>.

No special meanings, other than those attributed to the periphrasis in Vulgar Latin, have been verified for early Portuguese (from 1200 to 1500). Some authors have observed that lexical items apart from pronouns were infixes into the synthetic form, as in *Qu’eu apartar-me não hei* ‘For I’ll not be separated’ (as cited by Williams 1968: 207), which would provide evidence in favour of a tenuous fusion of this former periphrasis. However, the postposition of pronouns is found in the manuscript *Demanda do Santo Graal*, e.g. *...ou vós me matade, ou eu matarei-vos* ‘...either you kill me, or I’ll kill you’ (Melo 1957: 220). Beyond this, no specific comments are made regarding the awareness of the fusion

---

<sup>9</sup> In the ninth century two examples of the fusion of the infinitive + *habere* construction were already attested in French in the Strasbourg Oaths: *salvarai* and *prindrai*. The *Cancioneiros* mentioned by Maia (1986: 841) were written in Galician-Portuguese in the eleventh century, and the manuscript *Demanda do Santo Graal* is the Portuguese translation and adaptation of the French stories about King Arthur. It contains different versions of these stories and was written between the thirteenth and the fifteenth centuries.

nor any new meaning provided (Silveira Bueno 1967; Cardoso & Cunha 1978; Ferreiro 1996; Huber 1933; Maia 1986; Mattoso Câmara 1972; Nunes 1910; Said Ali 1921/ 2001; Silva Neto 1957/ 1986; Williams 1968).

### **2.3. Tracing the Meanings of the Future Forms in Portuguese Grammar Books**

In order to verify how grammarians treated the future forms in Portuguese over the centuries and, specifically, how they referred to the existence of the non-synthetic future forms, grammar books were examined from the sixteenth century onwards, namely from the century in which Portuguese grammar books were first published. The goal here was not to obtain historical explanations of the meanings of the future forms in the Portuguese language, but to explore the grammarians' synchronic viewpoint of the future forms.

This search was conducted in three stages: (i) grammars from the sixteenth to the eighteenth century were examined, (ii) grammars from the nineteenth centuries, and then (iii) grammars of the twentieth century. The investigation was designed to obtain access to as many grammar books as possible that were published during the period from the sixteenth to the nineteenth centuries. However, access to earlier Portuguese grammar books is restricted. Some original editions have not been reprinted and consequently were not available for consultation. Nevertheless, the most frequently cited grammar books were examined.

The selection of works from the twentieth century reflects the increased number of works published during this period. The samples were taken from: (i) grammar books traditionally mentioned in universities; (ii) compilations of letters with grammatical questions answered by grammarians; (iii) usage manual; and (iv) works for non-native

speakers of Portuguese. In order to follow the grammatical treatment of the future forms throughout this century, sources were selected from the beginning to the end of the century.

A total of 36 works were consulted from the sixteenth to the twentieth centuries.

Their distribution is as follows:

1500 – 1599	3 works
1600 – 1699	1 work
1700 – 1799	3 works
1800 – 1899	3 works
1900 – 2000	1 <sup>st</sup> half    9 works
	2 <sup>nd</sup> half    17 works

Table 2.1 displays the year of publication, the author, the classification<sup>10</sup> and the provenance of each work. PRT stands for Portugal<sup>11</sup>, BR stands for Brazil and O for other provenance. A complete bibliographic reference for each of these works is provided in Appendix 1.

---

<sup>10</sup> In the twentieth century, there is an increase in the number of meanings attached to the future forms by grammarians. It seems that there is an effort to describe the language; however, as explicated by Cunha & Cintra (1985/ 1999: xiv) “Trata-se de uma tentativa de descrição do português atual na sua forma culta, isto é da língua como a têm utilizado os escritores portugueses, brasileiros e africanos do Romantismo para cá,” ‘It is an attempt to describe the high-level Portuguese language, i.e. the language as it has been used by Portuguese, Brazilian and African writers since the Romantic Period.’ Thus, the grammars are based on the language of classical writers. As we will see in this study, they are still detached from spoken language, and still prescribe the non-interchangeable usage of the future forms. Based on this, in this dissertation I refer to the grammar books as *prescriptive*, unless other way stated (cf. Table 2.1).

<sup>11</sup> I noticed only one difference between Brazilian and European Portuguese regarding grammarians’ remarks. In the twentieth century (cf. Table 2.1), European Portuguese works pay more attention to the usage of the periphrastic variant *haver* + infinitive than Brazilian ones. A variationist study is needed to verify whether or not this is related to a high usage of this periphrasis in spoken European Portuguese.

**Table 2.1: Consulted Works**

<b>Year</b>	<b>Author</b>	<b>Classification</b>	<b>Provenance</b>
1536	Oliveyra	Grammar	PRT
1540	Barros	Grammar	PRT
1574	Gândavo	Grammar	PRT
1608	Leão	Grammar	PRT
1701	Justice	work for non-native speakers	O
1771	Reis Lobato	Grammar	PRT
1783	Melo Bacelar	Grammar	PRT
1809	Vieyra	Grammar	O
1852	Soares Barboza	Grammar	BR
1887	Silva Junior	Grammar	BR
1906	Figueiredo	Compilation of letters	PRT
1920	Amaral	Usage Manual	BR
1929	Figueiredo	Compilation of letters	PRT
1938	Machado Filho	Compilation of letters	BR
1941	Figueiredo	Compilation of letters	PRT
1944	Barreto	Prescriptive work	BR
1944	Barreto	Prescriptive work	BR
1944	Williams	work for non-native speakers	O
1945	Rossi	work for non-native speakers	O
1964	Corrêa	Grammar	BR
1964/ 69	Said Ali	Grammar	BR
1965	Nascentes	Descriptive work	BR
1968	Bechara	Grammar	BR
1968	Lapa	Grammar	BR
1968	Silveira Bueno	Grammar	BR
1969	Thomas	work for non-native speakers	O
1974	Thomas	work for non-native speakers	O
1983	Mateus, Brito, Duarte & Faria	Descriptive grammar	PRT
1984	Cunha & Cintra	Grammar	BR/PRT
1986	Kury	Prescriptive work	BR
1989	Kury	Prescriptive work	BR
1996	Van Achter, Monteiro, Teixeira & Duarte	work for non-native speakers	PRT
2000	Arruda	work for non-native speakers	PRT
2000	Azevedo	Descriptive grammar	BR
2000	Neves	Descriptive grammar	BR
2000	Sardinha & Ramos	Work on verbal conjugation	PRT

The search through these works revealed that the most relevant observations about the future in the Portuguese language are usually found in sections that deal with auxiliary verbs, special meanings and usage, in addition to the exceptions to the grammatical rules.

In general, grammarians consider the synthetic future to be the standard form of the verb conjugation, and hence is the form selected for verb conjugation tables. Grammarians usually depart from the distinction between time and aspectual/modal meanings in favour of establishing a distinction between the synthetic and periphrastic usages. They agree that the readings attributed to the periphrastic forms are basically the original meaning associated to their auxiliaries, which are totally or partially changed after their association to main verbs. Thus, the periphrastic forms are understood to hold special meanings, while the synthetic future is mainly assumed to express temporal reference. Moreover, the periphrastic forms in Portuguese, *haver* + infinitive and *ir* + infinitive, use their auxiliaries in the present tense, and together with the usage of the present verb form in future temporal contexts, are mainly identified with *certainty*, *determination*, *intention* and *proximate future*.

I summarized the results of my search in Table 2.2. This Table displays the meanings that are attached to the future forms by grammarians. The aim of this Table is to illustrate the treatment given to the forms in regards to their futurity meaning. When the grammarians cited the forms as denoting pure future time reference, I marked them for *futurity*. Each year either represents only one grammar, as in the case of the oldest ones, or more than one, as in the case of the twentieth century grammars.

*Table 2.2: Summary of the meanings attached to the future forms in usage manual, prescriptive and descriptive grammars of Portuguese by year (16th to 20th century).*

Meanings	1536	1540	1574	1608	1701	1771	1783	1809	1852	1887	1920	1941	1944	1945	1964	1965	1968	1969	1974	1983	1984	1986	1989	1996	2000
<b>Synthetic future</b> <i>Futurity</i> Positive statement Underdetermined Dubitative Certainty Imperative, suggestive Uncertainty, doubt Probability, possibility Politeness, questions Intention, determination, promise One syllable verbs	X	X	X	X	X	X	X	X	X X X	X	X	X	X	X X X	X X X	X	X X X X	X X X X	X X	X	X X X X X	X	X X X X	X X	X
<b>Have + infinitive</b> <i>Futurity</i> Latin participle future Obligation Intention, determination, volition, certainty Doubt, assumption, possibility, probability Necessity Proximity		X		X	X	X	X			X		X	X X X	X	X			X X X X	X		X X X	X	X	X	X
<b>Present</b> <i>Futurity</i> Certainty, necessary occurrence Proximity Emphasis, firm intention, promise High probability Unfeasibility, impossibility Politeness Adverbial specification								X		X		X						X X X	X X X	X X	X X	X	X	X	
<b>Ir + infinitive</b> <i>Futurity</i> Movement towards the accomplish of something Proximity Intention, resolution Certainty, necessary occurrence														X	X X X X	X	X	X X X X	X X X X	X X	X X X	X X	X X	X X	X

As Table 2.2 shows, the meanings attached to the Portuguese future forms are various. There is no consensus on them among grammarians. They often ascribe the same meaning to more than one future form, and the same form sometimes receives contradictory meanings, e.g., implications of both certainty and uncertainty are associated with the synthetic future form in the same grammar book (e.g. Cunha & Cintra 1984/ 1999: 457- 458).

This Table also shows that the synthetic future form is considered to be the future tense *per se* throughout the centuries. Only in the middle of the nineteenth and throughout the twentieth centuries do additional meanings begin to be attached to it. The other forms are mostly referred to as conveying specific meanings other than futurity since they were first mentioned.

From the eighteenth century onwards, the *haver* + infinitive periphrasis began to receive other meanings besides the correspondence with the Latin future participle, which included obligation. *Intention, determination, certainty, volition, possibility, necessity, assumption, probability, necessity, doubt and proximity* are all attached to its usage. Rarely, however, the *haver* + infinitive form is viewed as only expressing futurity and as being interchangeable with the synthetic future.

In general, the present variant follows the same path as the periphrastic *haver* + infinitive in the grammar books. It is acknowledged as expressing futurity from the nineteenth century onwards. However, throughout the twentieth century, other meanings were attached to its usage. These meanings are mainly equal to these attributed to the *haver* + infinitive periphrasis: *certainty, necessary occurrence, proximity, firm intention, high*

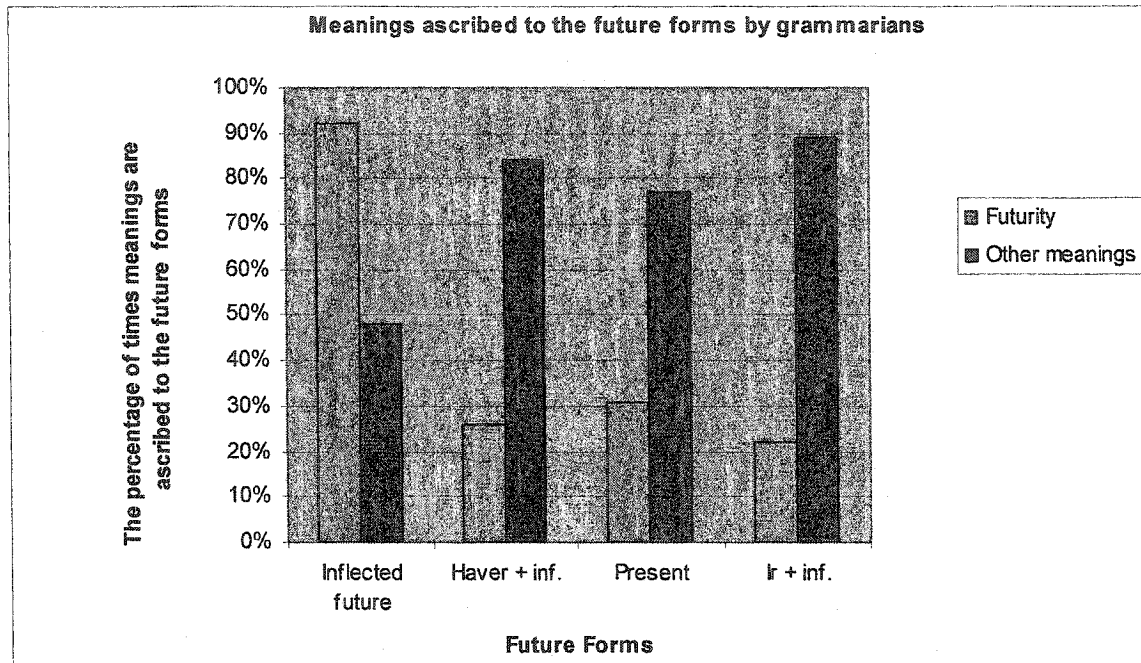
*probability*. Its usage is considered as an attempt by the speaker to frame the future event as real.

Only in the twentieth century did grammarians begin to cite the usage of the periphrastic form *ir* + infinitive in future contexts. It was first mentioned in a second language grammar without any additional meaning (Rossi, 1945); however, Portuguese grammarians did introduce this form later (cf. Table 2.2). They also gave this form other meanings. Besides the implication of movement, all the other meanings corresponded to those attributed to the present and the periphrastic form *haver* + infinitive. The grammarian's reasoning of the interpretation of this periphrasis seems to be the same as those of the interpretations of the usages of the present and the periphrastic form *haver* + infinitive: the speakers' choice to use the main verb or the auxiliary in the present tense indicates their engagement in the future action.

During the twentieth century, grammar books started to give more specific meanings to the synthetic form as well. However, at the end of the century, grammarians still referred to this form as expressing futurity. They commonly applied this meaning to the synthetic form as opposed to the others. Surprisingly, some of these works also acknowledged in notes or examples that the synthetic form was not used frequently in the spoken language. However, they did not offer any explanation for the apparent inconsistency between its exclusive status and its low frequency in speech.

The disproportion between the futurity meaning and the special meanings that are attached to the forms are easily visualized in the graph below. According to the summary in Table 2.2, the graph in Figure 2.1 displays the percentage of times each form is mentioned

as expressing futurity and the percentage of times each form is attributed other nuances of meanings.



**Figure 2.1: Percentage of Times Meanings are Ascribed to Future Forms**

Figure 2.1 shows that the periphrastic forms, i.e. *haver* + infinitive and *ir* + infinitive, and the present are mostly mentioned as conveying other meanings besides futurity. Their usages are mainly related to nuances of meanings which determine their precise selection in future temporal contexts. As this Figure shows, the reverse is valid to the synthetic future: this form is mentioned as expressing futurity 90% of the time, but less than 50% of the time it is attached to other meanings. Thus, the synthetic future is mostly remarked by grammarians as expressing futurity.

In Sections 2.3.1, 2.3.2 and 2.3.3, this grammatical research will be detailed with examples of grammarians' comments.

### *2.3.1. Future Forms in Grammar Books of the Sixteenth to the Eighteenth Centuries*

#### *2.3.1.1. Mentions of Variants*

I examined seven grammar books from the sixteenth to the eighteenth centuries: (i) Barros' (1540), Gândavo's (1574/1981), Olyveira's (1536); (ii) Leão's (1608/1983); and (iii) Justice's (1701/1970), Melo Bacelar's (1783/1996), and Reis Lobato's (1771). With the exception of Olyveira (1536) and Gândavo (1574), the other grammarians acknowledged the usage of the synthetic future and the periphrastic form *haver* + infinitive (cf. Table 2.2). The present and the periphrastic form *ir* + infinitive were not yet introduced in the grammar books.

#### *2.3.1.2. Mentions of Meanings*

The first two grammars of Portuguese were written by Fernão de Oliveyra and João de Barros in 1536 and 1540, respectively. Both authors, as well as Gândavo later on, attest the synthetic future form as the expression of the future (Barros 1540: 95; Gândavo 1574/1981: 27; Olyveira 1536: 72).

The periphrasis *haver* + infinitive is mentioned only by Barros (1540) in a section entitled *Dos Verbos Impessoaes* 'on non-personal verbs'. In this section, he states that this periphrastic construction denotes something to be done in a time that is still to come. He also states that it is used as a replacement for the Latin active future participle, which does not exist in the Portuguese language. Such an association is also made by historical

grammarians and linguists in the twentieth century for this periphrasis in early Portuguese (cf. section 2.2).

E quando se aiunta a uerbo e sempre e do módo infinito, e denóta algu auto por fazer: e per elle soprimos o participio futuro na uóz autiua, que os latinos te, de que careçemos, como. Eu ey deler os liuros, de que spero alcançár doutrina. (Barros 1540/ 1971: 93-94)

‘When it is linked to a verb which is always in the infinitive, it denotes that something is to be done. We use this form to translate the future active participle, which Latin people had, but we do not. For example: I shall read the books that I intend to master.’

In order to understand the exact meaning of *haver* + infinitive periphrasis in Portuguese as viewed by Barros, it was compared to the meaning of the Latin active future participle (Garcia 2000: 123). Basically, the active future participle corresponds to a relative clause with a verb in the future tense when translated into Portuguese, e.g. *uiri morituri* (*os homens que morrerão/ vão morrer* ‘the men who will die/ are going to die’). Hence, Barros assumes that *haver* + infinitive functions as the active future participle in Latin, the structure of which corresponds to a clause with a future-oriented verb. Thus, the usage of this periphrasis was not related to a semantic meaning, but rather to a syntactic structure.

In the seventeenth century, there is no change in the representation of the future forms. Leão (1608/ 1983: 298) mentions that in addition to the synthetic future form, which was formerly the infinitive of a verb pieced together with the Latin verb *habere*, the Latin future participle *amaturus* is expressed by *haver* + infinitive in the language; thus, *amaturus* in Portuguese would be: *o que há de amar* ‘one who shall love’. Other future

forms, such as the present and the periphrastic *ir* + infinitive, were not yet introduced in grammar books.

New meanings are associated to the usage of the periphrastic form *haver* + infinitive by grammarians in the eighteenth century (cf. Table 2.2). The first meanings are apprehended by Justice's examples (1701/ 1970: Dddd). This author states that the auxiliary verb *haver* in the present tense is added to an infinitive, as in (a): "Ey de amár, I am about to love, or I shall love; Ey de ler, ouvir, I am about to, or I shall read or hear, &c"; and also (b): "Sey que avéys de ser hómem dóuto, ou que serás hómem dóuto, I know that thou shalt or wilt be a learned Man". Justice not only presents the periphrastic future form in a separate section of his grammar, but also makes distinction when translates his examples. As cited in (a) and (b) above, the periphrastic form *haver* + infinitive is translated into English's forms *be about to* and *shall*, while the synthetic future is translated into the future form *will*<sup>12</sup>.

Melo Bacelar (1783/ 1996), however, clearly states that the word *obligation* is omitted in the periphrastic form *haver* + infinitive, but that this meaning is implicitly there; thus, one has to interpret the *haver* + infinitive construction as signifying that there is an obligation in the present time to perform an action in the future. This is the first time that a grammarian attaches the meaning of obligation to this periphrastic form:

---

<sup>12</sup> Justice (1701/ 1970: Eeee) contradicts the comparison between *haver de* + infinitive and *ter de* + infinitive constructions, in which the latter means obligation in the Portuguese language, when he states: "(...) Example, Ey de sér, amár, ler, I shall be, love, read; (...). The Verb Tenho in this Form is improper, altho' elegant enough in Spanish; for we do not say, Tenho de amár, ler, e[crever]; altho' the Spaniards say properly enough, tengo de amar, ler, e[scribir]".

Hei d'amar (h). (h) i.e. actualmente hei ou tenho obrigação d'amar em certo tempo futuro (...). Quem attender a que a nossa lingua está cheia d'abbreviações, ou supressoens pella figura Sillepsi, não duvidará subintender a dicta palavra obrigação. (Melo Bacelar 1783/ 1996: 56)

'Hei d'amar (h). (h) i.e. in the present time, I have an obligation to love in some future time (...). Those who know that our language is full of abbreviations, or suppressions by silepsis, will infer the word obligation in the periphrasis without doubt.'

Reis Lobato (1771: 79, 103) labels *haver* + infinitive as a *compound imperfect future* due to the auxiliary function of *haver*, but does not explicitly attribute any nuance of meaning to the periphrasis. I considered that he was the first grammarian to present the periphrasis as expressing only futurity (cf. Table 2.2):

O futuro imperfeito composto do indicativo fórma-se das vozes do presente do indicativo do verbo *Haver*, e da particula *De* anteposta á voz do presente impessoal do infinitivo do verbo conjugado, como v.g. no futuro imperfeito do indicativo do verbo *Amar*, em que se diz: *Hei de amar, Has de amar*, etc. (Reis Lobato 1771: 103)

'The indicative compound imperfect future is formed by the present of the indicative of *Haver*, and the particle *De* placed before the infinitive of the main verb, as in the case of the indicative imperfect future of *Amar*, one says: *Hei de amar, Has de amar*, etc.'

### 2.3.2. *Future Forms in Grammar Books of the Nineteenth Century*

#### 2.3.2.1. *Mentions of Variants*

As presented in Table 2.2, in the nineteenth century three future forms were acknowledged in Portuguese: the synthetic future, the periphrastic form *haver* + infinitive and the present. This century saw the first mentions to the present form usage in future temporal reference contexts (Silva Junior 1887 and Vieyra 1809).

#### 2.3.2.2. *Mentions of Meanings*

Three grammar books were perused: Silva Junior's (1887), Soares Barboza's (1852) and Vieyra's (1809). All three authors associated the synthetic future to *futurity*. As depicted in Table 2.2; however, in this century the synthetic future was for the first time interpreted as expressing other meaning than that of pure time reference: Soares Barboza (1852: 62) claimed that the synthetic future expresses a *positive*, but *undetermined* future time.

Futuro Absoluto (1). Sing. Eu louvarei, Tu louvarás (...) (1) É indeterminado, mas exprime de uma maneira positiva. (Soares Barboza 1852: 62)

'Absolute Future (1). Sing. I will praise, You will praise (...) (1) It is undetermined, but expresses a positive way.'

The periphrastic form *haver* + infinitive was given new meanings, namely, those of *possibility*, *firm resolution*, *necessity* and *intention* were mentioned (Vieyra 1809; Soares Barboza 1852). This periphrasis was considered, for instance, as expressing a present intention to perform a future action, *in verbis*:

1. O *Infinitivo Impessoal* exprime a afirmação desacompanhada das circunstâncias de Pessoas e de Tempos, ou seja a afirmação não acabada, como *Louvar*, ou acabada, como *Ter louvado*, ou principiada na tenção, e futura na execução, como *Haver de louvar*. (Soares Barboza 1852: 67) (emphasis added)

'1. Impersonal infinitives express actions without Person and Time circumstances, i.e. unfinished actions, as *Louvar* 'to praise', finished actions, as *Ter Louvado* 'to have praised', or actions that were started in their intention, but were future in their execution, as *Haver de louvar* 'to have [of] praise'.

As mentioned in the previous section, the present form is already regarded as a future exponent at the beginning of this century (Vieyra 1809). Table 2.2 shows that, although Vieyra did not claim any special meaning for the usage of the present form in future contexts, Silva Junior (1887) stated that in such cases the future event or action is *proximal or connected to the present*.

In summary, by the end of the nineteenth century, grammarians have considered three forms as potential expressions of the future in the Portuguese language: the synthetic form, the periphrastic form *haver* + infinitive and the present. Still, the synthetic future form is the only one that officially represents the future on verb conjugation tables and the periphrastic *haver* + infinitive is referred to at first as merely a construction replacing a flexional Latin construction, which is translated into Portuguese by a relative clause with its verb in the future. It is later in the eighteenth century that this construction starts to be attributed with more meanings (cf. Table 2.2). From the sixteenth to the eighteenth centuries, the grammar books do not refer to the present form as expressing future. This is first mentioned in the nineteenth century by Vieyra (1809) and Silva Junior (1887). Until the end of the nineteenth century, the periphrastic form *ir* + infinitive had not been

considered as expressing a future time in Portuguese. This periphrastic construction was only cited by Justice (1701/ 1970: Dddd)<sup>13</sup> as being used in the past tense.

### 2.3.3. *Future Forms in Works of the Twentieth Century*

In the twentieth century, there is evidence in prescriptive literature that all four variants of the future temporal reference were established in the language. Thus, it was expected that the grammar books of this period be concerned with the usage of these forms in the language. However, despite a considerable growth in the number of examples with future form usages, with a few exceptions, the grammarians were more concerned with the interpretations of the future form usages than their actual usage in the spoken language.

As the doctrine of language correctness was widely implemented in the twentieth century, a large number of grammatical works were published, although they mainly reiterated earlier grammatical statements. For the purpose of this investigation, the works of well-known grammarians were selected (Amaral 1920; Bechara 1968; Corrêa 1964; Cunha & Cintra 1984/ 1999; Kury 1986; Kury 1989; Lapa 1968; Mateus et al. 1983; Nascentes 1965; Said Ali 1964, 1969; Silveira Bueno 1968). I also searched for comments regarding the future temporal reference in letters answered by grammarians (Figueiredo 1906, 1929, 1941; Machado Filho 1938; Barreto 1944, 1944/ 1980), in Brazilian and European Portuguese second language books (Arruda 2000; Rossi 1945; Sardinha &

---

<sup>13</sup> Justice (1701/ 1970: Dddd) gives an example of the usage of the periphrastic *ir* + infinitive in the past when he explains the difference between the verb to be and the verb to go in the preterperfect tense. He states: "The verb *Sou*, I am, is exactly expressed and pronounced in the Preterperfect Tense as the Verb *Vou*, I go, by the word *fuy* (...); only with this Distinction, that the Preterperfect Tense of the Verb *Vou*, I go, is always followed by a Substantive of Place and Motion (...) Sometimes follows a Verb of the Infinitive Mood; as *Fuy* ver, I went to see; *Fuy* ler, I went to read, &c".

Ramos 2000; Thomas 1969, 1974; Van Achter et al. 1996; William 1944), and in two recently published grammar books (Azevedo 2000; Neves 2000) (cf. Table 2.1).

### 2.3.3.1. *The Synthetic Form*

Grammarians continue to represent the future tense in verb conjugation tables by means of the synthetic future form. However, during the twentieth century, and presently, the synthetic future has been more commonly associated with other meanings besides futurity. Many grammarians claim that this form conveys *uncertainty, doubt, probability* and *possibility* (Arruda 2000: 135; Azevedo 2000: 167; Bechara 1968: 338; Cunha & Cintra 1984/ 1999: 457-58; Kury 1989: 135; Rossi 1945: 203; Said Ali 1964: 319, 1969: 164; Thomas 1974: 123; Van Achter et al. 1996: 155). The use of the synthetic future is also seen to imply that the speaker is *certain* of the realization of a future event or action (Cunha & Cintra 1984/ 1999: 457; Kury 1989: 134).

In addition to these meanings, the synthetic future is said to be used for the *imperative* (Arruda 2000: 136; Bechara 1968: 339; Cunha & Cintra 1984/ 1999: 458 and 477; Rossi 1945: 203; Said Ali 1964: 317, 1969: 164; Silveira Bueno 1968: 345; Van Achter 1996: 151), which is explained as a result of the absence of the Latin imperative future form in the Portuguese language (Kury 1989: 135). *Politeness* (Arruda 2000: 136; Bechara 1968: 338; Cunha & Cintra 1984/ 1999: 457; Said Ali 1969: 165), *intention, determination* to perform an action and *promise* (Mateus et al. 1983: 173; Thomas 1969: 123; Van Achter et al. 1996: 155) are new meanings attached to this form during this period. Amaral (1920) advocates that its usage is limited to *dubitative* contexts, while

Thomas (1969: 123) suggests that it is mainly restricted to *monosyllabic verbal stems* in the language.

In the second half of the twentieth century, a few grammarians provided evidence of the infrequent usage of the synthetic future in the spoken language. They made notes stating that this form was *relatively rare, hardly ever used, used comparatively little* in spoken Portuguese, and *used essentially in the written language* (Cunha & Cintra 1984/1999: 459; Kury 1989: 136; Thomas 1969: 123; 1974: 117; Van Achter et al. 1996: 155).

Na língua falada o futuro simples é de emprego relativamente raro. (...) (Cunha & Cintra 1984/1999: 459)

‘In spoken language, the usage of the simple future is relatively rare.’

Notwithstanding, these same grammarians continued to ascribe a list of meanings to the synthetic future as opposed to the other forms, as mentioned in the previous paragraphs (cf. Table 2.2.).

#### 2.3.3.2. *Haver + Infinitive*

Prescriptive works focused on the synthetic future and the periphrastic form *haver + infinitive* in the first half of the twentieth century. Their main concern was on the placement of clitics with these future forms. Since, in Portuguese, clitics must be infixated into the synthetic future, this issue was often discussed by prescriptive authors (Figueiredo 1906, 1929, 1941; Machado Filho 1938; Barreto 1944, 1944/1980), as exemplified in (1) and (2) below:

- (1) “Com as formas compostas, a construção mais comum é *há de responder-lhe*, por onde *responder-lhe há de*, *responder-lhe há*. (...) Sómente do século XIV em diante, como informa Said Ali, começa a vogar a construação do tipo *responder-lhe-á*, em que se opera a chamada mesóclise. Dantes surgiam destas belezas: *tornarei-me entam* (S.Graal, 18)...” (Machado Filho 1938: 111-112)

‘The most common construction with compound forms is *há de responder-lhe*, corresponding to *responder-lhe há de*, *responder-lhe há*. (...) According to Said Ali, this construction starts to be used with mesoclysis - *responder-lhe-á* – only from the XIV century onwards. Before this century, there were strange things like: *tornarei-me entam* (S. Graal, 18)...’

- (2) “Ora, o futuro simples é *darei*; não é verdade? Pois êste futuro, que é a simples inversão de *hei dar* (ou *hei de dar*), deveria produzir, segundo os princípios da minha amabilíssima correspondente, *dar + hei = darhei*. (...) Conjugado êste tempo pronominalmente, temos *te darei* e *darei-te*. Esta segunda forma, porém, não se usa, e, por convenção, ou por eufonia, ou pela figura chamada *tmese*, o *te* intercalou-se no *darei*, produzindo necessariamente *dar + te + ei = dar-te-ei*.” (Figueiredo 1941: 93)

‘Well, the simple future is *darei*; isn’t that right? This future, which is merely the inversion of *hei dar* (or *hei de dar*), should produce, according to my correspondent, *dar + hei = darhei*. (...) when used with a pronoun, it becomes *te darei* e *darei-te*. The latter form, however, is not used. For convention or euphony, or for *tmese*, the pronoun is placed in the future form *darei*, from which *dar + te + ei = dar-te-ei*.’

It is worth noting that when authors discuss the prescriptive rule of clitic placement with the future, they did not attach any specific meaning to the periphrastic form *haver + infinitive*. On the contrary, they show both forms as equally representing the future: *haver (de) + infinitive* or *infinitive + haver* (cf. excerpts (1) and (2) above). Regardless of this, grammarians have assigned a list of interpretations to the periphrastic form *haver + infinitive*.

The most common source of prescriptive comments regarding the usage of the periphrastic form *haver* + infinitive is the correlation of this periphrasis to the Latin future participle structure (cf. section 2.3.1). Based on this comparison, grammarians usually associated the meanings of *necessity* and *obligation* to this periphrastic form. In the twentieth century, grammarians still interpreted the usage of this periphrasis as the speaker's wish to convey feelings of necessity or obligation of performing a future action (Barreto 1944/ 1980: 513; Bechara 1968: 135; Cunha & Cintra 1984/ 1999: 393 and 532; Said Ali 1969: 73; Silveira Bueno 1968: 156; Rossi 1945: 203; Thomas 1969: 124, 1974: 117). Moreover, a correspondence of this form to "the strong English future *will* in contrast to *shall*" is explicitly made by Rossi (1945: 91).

Nenhuma novidade existe no uso dêste tempo que continua a ser o mesmo, hoje, qual fôra no tempo antigo. (...) Em forma composta haverá o significado de absoluta obrigação. (Silveira Bueno 1968: 156) (emphasis added)

'There is no innovation in the usage of this tense that is still the same today as it was before. (...) There will be the meaning of total obligation for the compound form.'

Many other meanings are connected to this form: (i) the *intention* of doing an action in the future (Arruda 2000: 136; Corrêa 1964: 31; Cunha & Cintra 1984/ 1999: 459; Kury 1989: 136; Sardinha & Ramos 2000: 142; Thomas 1969: 124); (ii) *probability*, *possibility*, *conjecture* or *supposition* (Thomas 1969: 124; Barreto 1944: 148; Williams 1944: 70); and (iii) *volition* (Kury 1986: 43). Furthermore, an implication of *promise*, *determination* and *certainty* are also associated with it (Corrêa 1964: 31; Kury 1989: 136; Said Ali 1964: 319; Sardinha & Ramos 2000: 142; Thomas 1969: 124).

Although prescriptive literature attaches several readings to the usage of *haver* + infinitive, only Thomas (1974: 117) mentions it as a means of expressing only pure futurity: “this form may have the meaning ‘He is to speak, he is supposed to speak, etc.’, but it is often exactly equivalent to the simple future”. Surprisingly, though rarely, this form is regarded as being more frequently used than the synthetic form in future temporal contexts (Said Ali 1964: 319; Thomas 1974: 117).

### 2.3.3.3. *The Present Form*

The present form in future temporal reference contexts is well acknowledged in the twentieth century grammars; however, its reference to future temporal actions and events is associated with *certainty* rather than with futurity (Cunha & Cintra 1984/ 1999: 449; Kury 1989: 130; Lapa 1968:157; Nascentes 1965: 167; Said Ali 1964: 311; Thomas 1969: 116; Van Achter 1996: 167). In addition to certainty, the present form is mentioned as conveying *firm intention, emphasis, promise* (Bechara 1968: 336; Kury 1989: 136; Said Ali 1964: 162; Thomas 1969: 116), *necessary occurrence, high probability, unfeasibility, impossibility* (Mateus et al. 1983: 121), *politeness* (Said Ali 1969: 162) and *proximity* (Cunha & Cintra 1984/ 1999: 448; Said Ali 1969: 161; Silveira Bueno 1968: 343).

Examples:

Emprega-se o presente: (...) (b) pelo futuro do indicativo para indicar com ênfase uma decisão: Amanhã eu vou à cidade. (Bechara 1968: 336) (emphasis added)

‘Use the present: (...) (b) for the future to indicate an emphatic decision: Tomorrow I go to the city.’

Muitas vezes emprega-se o tempo presente para designar o ato inexistente, mas de realização próxima. (Said Ali 1969: 161) (emphasis added)

‘Many times the present tense is used to designate an inexistant action, which will be accomplished soon.’

With the exception of the meanings of unfeasibility, impossibility and politeness, all the other meanings result from the reasoning that speakers frame future actions or events in the present tense as a consequence of their emotional attachment to that event. Thus, prescriptive literature advocates this usage as illogical and as a manner of transforming uncertain facts into reality, since they are not yet accomplished. We can witness such reasoning explicitly when grammarians explain the usage of this form. For instance, the correctness of this usage was questioned in a letter addressed to a grammarian who asserted that it was logically wrong (Barreto 1944: 249-250). Lapa (1968: 157) exemplifies this process of bringing a future event to the present time:

“(...) ‘Amanhã *chega* o teu primo José; *vais* à estação esperá-lo’. A fantasia e o sentimento aproximam de nós o fato futuro e incerto; a fantasia torna-o presente, o sentimento torna-o coisa certa (...). Lògicamente, aquela frase deveria ser assim redigida: ‘Amanhã *chegar*á o teu primo José; *irás* à estação esperá-lo’.”

‘(...) Tomorrow your cousin José *arrives*; you *go* to the station to wait for him. Fantasy and emotion bring an uncertain and future event nearer to us; the fantasy converts it into present, and the emotion gives it certainty (...). Logically, that sentence should be written like this: Tomorrow your cousin José *will arrive*; you *will go* to the station to wait for him.’

Therefore, the present is recognized by grammarians as being used in future contexts; but it is not prescribed to express futurity.

Another comment regarding the usage of this form in the spoken language is the necessity of an adverb of time in the utterances. Since, as it is explained, basically all references to future states or events by the present form need to be clearly situated in the future. In other words, adverbs of time, rather than verbs themselves, become responsible for disambiguating the temporal reference of utterances (e.g. Cunha & Cintra 1984/ 1999: 438; Thomas 1969: 116; Van Achter et al. 1996: 155):

*Instead of the future tense, to express future action. The simple present may be used whenever the time of action is made clear by an adverb of time, by any other expression of time, or by the context. (Thomas 1969: 116) (emphasis added)*

Amaral (1920: 55) and Thomas (1969: 116) state that the present form is the one used most frequently in the future temporal context in Brazilian Portuguese. Further in his book, Thomas (1969) claims that the present form is widely accepted in such a context because the present tense is expressed by the progressive form in the language; thus, speakers do not interpret the simple form of the present as being used to express present time:

The most common of all forms used to express future time in BF is the simple present. The development of the progressive form in Brazil to the point where its differentiation from the simple present is obligatory reduces materially the possibility of confusion when the present is used for the future. (Thomas 1969:126)

[BF = Brazilian Future]

#### 2.3.3.4. *Ir + Infinitive*

The last future form traced in the Portuguese grammars was *ir + infinitive*. The grammarians assign a meaning of *movement* to this form primarily carried by the auxiliary; as a result, the periphrasis is interpreted as a *movement* towards the accomplishment of an act (Bechara 1968: 136; Thomas 1969: 125; Said Ali 1964: 174). Then, the meanings of firm *resolution, intention, certainty* and *necessity* of occurrence are applied to the form, (Cunha & Cintra 1984/ 1999: 395; Kury 1986: 42; Mateus et al. 1983: 120; Said Ali 1964: 174; Thomas 1969: 125; Van Achter et al. 1996: 155). Another recurring meaning attached to this periphrasis is the *proximity* of the event or the action (Cunha & Cintra 1984/ 1999: 459; Kury 1986: 42; Said Ali 1964: 174; Thomas 1969: 125; Arruda 2000: 136):

A combinação de *ir + infinitive* pode designar locomoção ou também simples intento de pôr alguma coisa em efeito, ou ainda, um fato que não tardará a realizar-se: *Vou tomar o trem.* (Said Ali 1964: 174)

‘The combination *ir + infinitive* may designate motion or intention to do something, as well, a fact that will be realized soon: *I’m going to take the train.*’

Auxiliares que determinam com mais rigor o momento do processo verbal, indicando, entre outros, os seguintes “aspectos”: (...) e momento futuro próximo: *ir + infinitivo*: ‘VOU viajar para a Catalunha.’ (Kury 1986: 42) (emphasis added)

‘Auxiliaries which determine the moment of the verbal process more precisely, indicating the following aspects, among others: (...) e proximate future moment: *go + infinitive*: ‘[I]’ m going to travel to Catalonia.’

This review demonstrates that grammarians ascribe to the periphrastic form *ir* + infinitive the same nuances commonly attributed to the periphrastic form *haver* + infinitive and the present form. With the exception of movement, all the other meanings are based on the fact that the auxiliary *ir* is morphologically conjugated in the present tense. Hence, speakers bring to the present time an uncertain future eventuality; consequently, they give it a colour of reality. The underlying reasoning that grammarians follow here is the same used to explain the *present* and the *haver* + infinitive usage by means of these meanings.

Although *ir* + infinitive is typically seen as being used only when speakers mean something other than futurity, some grammarians do not attribute any special meaning to this periphrasis. They simply ascribe futurity to its usage (Neves 2000: 65; Rossi 1945: 99; Sardinha & Ramos 2000: 143). Remarks regarding its high frequency of usage in the spoken language are made by Azevedo (2000: 167) and Thomas (1969: 125).

#### 2.4. Summary

The previous sections focused on both the acknowledgment and the meanings that the Portuguese future forms have received throughout the centuries. It should be noted that the analyses of the future forms were based mostly on the explanations that favoured their differences in meanings rather than their interchangeable usage in the language, even though some remarks showed that the latter was the case.

For instance, both historical and prescriptive grammarians agree that *hei* + infinitive and infinitive + *hei* corresponded to the same Latin participle future, which was interpreted as a future-oriented form in Portuguese. Their meanings were the same. The periphrastic form infinitive + *hei* underwent agglutination and consequently was chosen to be the verbal paradigm. This piece of information may lead us to infer that this form was already

grammaticalized, while *hei* + infinitive was used with other meanings or was not used to express futurity as frequently as infinitive + *hei* in the language.

However, some historical remarks (cf. section 2.2., Nunes 1910: 320; Williams 1968: 210) show that this may not be the case. In early Portuguese, infinitive + *hei* and *hei* + infinitive were in variation. The written tradition elected for the former, but even in early texts lexical items and pronouns were inserted between the infinitive + *hei*, which shows that the fusion was not so solid. On the other hand, the written tradition inserted a preposition between *hei* + infinitive, even though it was used as a unit (*hei* infinitive). The first grammar books selected infinitive + *hei* to the verbal paradigm because it conformed to an inflected pattern, which mirrored the Classical Latin inflected paradigm; i.e. suffixes attached to infinitives. It seems that it was selected as the future form because only one form is adequate to represent one tense, i.e. form – function symmetry.

From this point of view, we see that the grammatical tradition, which has succeeded mostly due to the written tradition, has promoted the synthetic future as the expression of futurity, and discouraged the other future forms, when acknowledged, as illogical or as conveying numerous other meanings, which require their careful usage in future contexts.

The search for the meanings attributed to the future forms shows that from the very first grammar books, grammarians have not invested in the effort to represent the spoken language. Rather, they selected and prescribed one form for one function. When they invoke new forms related to the same function, each form is ascribed a nuance of meaning. Thus, they can maintain the appearance of form/function symmetry. This process clearly demonstrates that there is a refusal of the fact that more than one form may represent a same function in language (cf. section 3.1.1), a rejection of inherent variability.

As we will see in chapters 5 and 6, the actual use of the synthetic future is far from the way grammarians have conceived it to be in Portuguese. Since the time of the first Portuguese grammar books, the future forms have been competing for the future context in the language. The variationist methodology will enable the grammarians' comments regarding the future forms to be tested against their real usage.

## **2.5. Future Form Meanings in Prescriptive Works and their Handling by Variationist Research**

The fact that grammarians acknowledge the existence of more than one future form does not mean that they accept the form-function asymmetry. On the contrary, as has been revealed, they establish a relation between form and function in which every difference in form must be conceived as a result of a difference in meaning or function, however tenuous (cf. section 3.1.1). Furthermore, they do not work only with future temporal reference contexts. Thus, many meanings that are attached to future forms do not represent the future, for instance, the imperative, doubt, probability, supposition meanings that are conferred to the synthetic future form.

The miscellaneous interpretations confer subtle differences to the forms. These subtle differences are many times impossible to test because the future form usages are associated, for instance, with speakers' mental picture of a situation. Variationist studies have provided evidence that there is little support for many semantic interpretations associated to future variants, and also that in spontaneous discourse, a speaker's intention and a hearer's inference are not reachable. The access we have to speakers' intentions is provided through their speech in a natural situation only (Sankoff 1988a; Poplack 1992; Poplack & Turpin 1999; Poplack & Tagliamonte 2000; Poplack & Tagliamonte 2001).

The variationist procedure is similar to panning for gold. Although the grammarians' remarks display a great deal of unsystematic correlations, a number of possible constraints can be extracted from their work, operationalized as factors and empirically tested against the data by means of the quantitative methodology. These factors will be presented in chapter 4 of this dissertation.

## **Chapter 3** *Theoretical Framework, Method and Data*

---

This chapter presents the theoretical framework, methodology and data used in this investigation. The focus is on the spoken language and the inherent variability that characterizes it.

### **3.1. Theoretical Framework and Methodology**

The type of research undertaken in this dissertation is known as Quantitative Sociolinguistics or the Theory of Variation. The model followed here is the one proposed by Weinreich, Labov and Herzog (1968), in which language is considered to be an inherently variable heterogeneous structure:

Long before predictive theories of language change can be attempted, it will be necessary to learn to see language – whether from a diachronic or a synchronic vantage – as an object possessing orderly heterogeneity. (Weinreich et al. 1968:100)

The theory of language outlined by the above authors is conceived as part of an investigation in which the ultimate goal is to determine the course of evolution of present-day languages. Thus, the investigation should establish the direction of the linguistic evolution and the factors that affect this evolution. Such theory deals with basically four different temporal ranges: long-term changes, completed changes, ongoing changes and synchronic sections of a change. The aforementioned authors establish five features that must be addressed in the investigation of language variation and change: (i) the constraint,

or the possible changes and their conditions; (ii) the transition, or intervening stages of a change; (iii) the embedding, or social and linguistic motivation of a change; (iv) the evaluation, or level of social awareness in the linguistic change; and finally (v) the actuation, or the ability to relate the behaviour of a single variable to the whole system of a specific language (see also Labov 1972a: 161-2). Attention to these five stages will allow the investigator to model the variable use of linguistic structures in a specific language.

While some linguists rely on their native or intuitive knowledge of the linguistic system of a specific language to provide data for their analyses, Variationist Theory asserts that actual spoken language is the only reliable source of data for studying variation. In order to systematically account for the facts of language, researchers need to deal with the variability within the system. No invented or isolated datum is able to cover that variability. Therefore, variationist analyses are based on a quantitative investigation of a variable rule, i.e. a particular grammatical function expressed by two or more alternative forms (variants), within a large body of spontaneous spoken data. The first empirical study within this framework was carried out by Labov (1966).

Due to the fact that the choice process of a variant is influenced by different factors that are differently combined in contexts, and since investigations on variable phenomena are necessarily anchored in empirical evidence, this type of research can hardly be done without a statistical program of multivariate analysis. The basic procedure of the statistical program is explained in detail in Sankoff (1988b; see also Naro 1981; Naro et al. 1999, Scherre & Naro 1999). This work makes use of the statistical package program GoldVarb 2.0., which was designed for Macintosh (Rand and Sankoff 1990).

As posited by Tagliamonte (2002: 731), 'One of the foundations of variationist analysis is its attempt to discover not individual occurrences or overall rates of occurrences,

but *patterns* of variability in the body (or bodies) of material under investigation'. Using the multivariate analysis, researchers are able to uncover the pattern of variability.

Researchers follow a principled series of steps in order to discern this pattern. First, it is necessary to circumscribe the variable context, i.e. the linguistic context in which alternate forms occur interchangeably. Second, the data are coded into factors which represent working hypotheses about the choice of those alternate forms. Then, the multivariate analyses reveal the independent factors that favour or disfavour a given variant and the strength of their effect. This statistical method assesses the contribution of each of the different factors combined in the context to the choice of the variants in three ways (cf. Poplack and Tagliamonte 2001: 94; Tagliamonte 2002: 733; Van Herk 2002: 48):

(i) *statistical significance* of the effect ( $p = <.05$ ) shows whether the effect of factors is greater than a random effect. Researchers aim to ascertain which factors are statistically significant to the realization of variants;

(ii) *magnitude of effect*, as expressed by the *range*, i.e. the difference between the contributions of the most and least favouring factors within each group of factors, shows which of the selected group of factors influences the most the choice of variants. Researchers aim to find out the group of factors that are most or least significant to the variation;

(iii) *constraint hierarchy* shows whether the factors constrain the variant in the direction expected by the hypotheses. It is the ordering of probabilities within a group of factors. Even if not significant, the direction of effect may support or undermine a researcher's hypothesis. It may also be used to trace the path of linguistic development of a variable by comparing the constraint hierarchy of group of factors throughout the time.

The multivariate analysis enables the researcher to uncover the underlying grammatical structure of a variable. More important for the study presented here, it will allow the researcher to establish the factors that favour each future temporal referent throughout the centuries and to identify the contexts in which the usage of these referents expands or retreats in the language.

Conceiving of language as it has been presented in this section implies that researchers accept two assumptions: first, that variation presupposes a non deterministic form-function relationship in languages; second, that the only plausible way of scientifically studying a language is through real production; thus, the data cannot be fabricated nor be represented by extractions of isolated examples without analyses of its regularity in natural production.

### *3.1.1. Form – Function Asymmetry*

One fundamental issue regarding linguistic variation is the characterization of the form-function relation in language. The concept which has traditionally been accepted in linguistics is the unique relation of one grammatical function to one form, as claimed by Bolinger (1977: x, 1): “the natural condition of a language is to preserve one form for one meaning, and one meaning for one form” and “if two ways of saying something differ in their words or their arrangement they will also differ in meaning...”. This view, which underlies many linguistic theories, is contradicted by the undeniable fact that all languages evolve, and that changes over time are always a result of synchronic variation, i.e. a speaker’s choice between different forms of expression of the same function.

Therefore, in contrast to the one form - one function symmetric relation, natural speech provides examples in favour of an asymmetric relation between forms and function, as shown in (1) with data from the Portuguese Future Temporal sector:

- (1a) Velhaco insolente, tantas me tem feito, que agora te *mandarei* (SF) enforçar. (Silva 1734. VE: Rei: second part: scene 8: 145)

‘Impertinent man, you have caused me so much trouble that I *will order* your hanging.’

- (1b) Prometo tudo; pois vejo que tu não *hás de fazer* (HP) a torre no ar, e assim sempre te venho a matar; vamo-nos e levem-no preso para que não fuja. (Silva 1734. VE: Rei: second part. scene 8: 145)

‘I promise you everything because I know that you *will not build* the tower in air. Thus, I will kill you; Let us go. Arrest him, so that he does not run away.’

This form-function asymmetry can also be corroborated in situations where one form can express more than one function. For instance, verbs in the present tense are used to express both a present temporal reference and a future temporal reference in Portuguese, as in (2):

- (2) *São* nove horas, e o ofício de Ramos *principia* (P) às dez e meia. (Pena 1853. N: Ambrósio: act 1: scene 2: 12)

‘It *is* nine o’clock, and the Palm procession *will begin* at half past ten.’

The variationist framework is appropriate for investigating systematically the contexts in which competing forms occur interchangeably, as in (1 a, b), including those in which the differences among the forms, if they exist, are neutralized (Sankoff 1988a). The

resistance of languages to strict form-function symmetry explains many of the misconceptions about language structure and language usage.

As was seen in chapter 2, the explanations of many grammarians regarding the use of the future forms in Portuguese are based on subtle semantic distinctions that are imposed on the future forms as an attempt to describe logically their different usage in the Portuguese language. However, once a corpus of natural discourse is examined, we notice that the different future forms co-occur in the same future temporal reference context, i.e. they have exactly the same function. Thus, the form-function asymmetry cannot be reduced or eliminated. Indeed, it must be treated as an inherent property of language.

### **3.2. The Scope of the Investigation**

In the analysis of variation in the future temporal reference context in Portuguese, a motivating issue is the examination of the course of variation/ change in the usage of the future forms throughout the centuries. In order to trace this evolution, this work takes into account diachronic and synchronic data from Portuguese.

The diachronic component is inspired by preliminary observations of the incongruity between the usage of future variants in spoken Brazilian Portuguese and the prescription of future forms in grammar books (cf. chapter 2). Since variation in the future context is not stigmatized in Brazilian Portuguese, this inconsistency is inferred to be a result of linguistic change in this sector, in addition to the inability of grammarians to accept language variability. For this reason, trajectory of the variants is traced over the centuries. The usage of variants is investigated from the sixteenth to the nineteenth centuries in written language. The constitution of the data set is detailed in the next section.

The synchronic component reveals the current usage of future variants in Brazilian Portuguese. This component includes plays written in various years in the twentieth century, as well as speech from the end of that century. The data from the twentieth century plays make it possible to trace the trajectories of the variants and the factor effects in a corpus that shares the same characteristics of those that will be examined in the diachronic scope, i.e. they represent the spoken language. In addition, the twentieth century speech data enable me to attest the conditioning factors that constrain this variable in spoken Brazilian Portuguese. Furthermore, the results obtained from speech are vital for validating the written data as being representative of the spoken language.

### **3.3. The Data Set**

#### *3.3.1. The Diachronic Data*

The study of the structure of languages and their past features is usually based on the reconstruction process only. The transition from one stage of the language to another is commonly ignored and competing structures unknown. The theory of variation opens up a new way of investigating the recent past of a language without ignoring its changing courses. By taking into account a population's successive age levels, quantitative analysis is able to indicate any significant change in the use of variants over a period of time. Thus, using data from the present and the concept of "apparent time", researchers can find out the trajectory of variation.

Although the present can be used to illuminate the past in cases of recent change, diachronic data<sup>14</sup> is needed to explain language variation and/ or change that occurred in two or more centuries ago. The attempt to obtain the vernacular of centuries ago by means

---

<sup>14</sup> Unless it is possible to replicate variationist studies using data from language that is considered to be representative of an early stage of other spoken language.

of documents and literary texts is not assured. Historical documents were usually produced in special contexts and tend to be very formal or formulaic. Moreover, they do not survive with enough social information about the writers. Almost the same may be claimed about literary texts. The language is framed by a writer's style, a literary style, the influence of standardized language on the writers or even by the literacy level of the community.

Regarding the limitation of historical documents, Labov (1994: 11) observes the following:

Historical documents survive by chance, not by design, and the selection that is available is the product of an unpredictable series of historical accidents. The linguistic forms in such documents are often distinct from the vernacular of the writers, and instead reflect efforts to capture a normative dialect that never was any speaker's native language. As a result, many documents are riddled with the effects of hypercorrection, dialect mixture, and scribal error. Furthermore, historical documents can only provide positive evidence. Negative evidence about what is ungrammatical can only be inferred from obvious gaps in distribution, and when the surviving materials are fragmentary, these gaps are most likely the result of chance.

The restrictions imposed on the use of historical documents are also applicable to the use of literary sources, as observed below:

(...)We usually know very little about the social position of the writers, and not much more about the social structure of the community. Though we know what was written, we know nothing about what was understood, and we are in no position to perform controlled experiments on cross-dialectal comprehension. Our knowledge

of what was distinctive and what was not is severely limited, since we cannot use the knowledge of native speakers to differentiate nondistinctive from distinctive variants. (Labov, 1994: 11)

It is important to bear in mind that the most reliable data are collected by variationist techniques, such as the sociolinguistic interview. However, the attempt to understand the origin and the path of a long-term variable phenomenon leads us to the hard task of making use of the best available diachronic source.

The first known written Portuguese texts are from the twelfth century when Portugal was recognized politically as a nation. However, because the new kingdom was in the process of establishing itself as an independent country, these texts relate mainly to legal and religious issues. Hence, it is highly unlikely that the vernacular would be represented in these documents, especially since the new kingdom also wished to distinguish its language from Galician-Portuguese and Spanish and so decided to maintain Latin's influence as proof of its cultural dominance.

It is possible to draw an analogy between the first three centuries of the Portuguese language in the new kingdom of Portugal and those of the Brazilian Portuguese language in Brazil. In the history of both countries, the process of recognition and acceptance of the language followed the political and economic performance of the nation in relation to their dominant country, i.e. Spain in the case of Portugal and Portugal in the case of Brazil. Over the first few centuries, all documents and literary texts were written in European Portuguese in Brazil. With the exception of poetry, Brazilian Portuguese was not recognized as a literary language until the early nineteenth century. Thus, it is highly unlikely that any vernacular Brazilian Portuguese will be found in literary production before then.

A knowledge of history, history of language, demography, and literature as well as the social organization of communities is necessary to be able to search for reliable written data. Considering that historical documents are rather limited in their ability to show the regularity of variation at a specific time, it was decided to refer to plays. These are by far the best diachronic sources of speech that we have when they contain colloquial situations, colloquial characters, representations of different social classes, and vernacular linguistic features<sup>15</sup>. Though Labov's remarks are also applied to the literary production; popular plays, and more specifically comedies, farces and comedies of manners<sup>16</sup>, mirror the spoken language in a more reliable way, as the next section will show.

### 3.3.1.1. *The Selection of Plays*

#### 3.3.1.1.1. *Portuguese Plays*

In the Middle Ages, the Church had the power to determine the content of theatrical production in Portugal. Christian rituals and Jesus' story were originally acted out only in churches, but later on in castles and public places as well. The so-called profane performances, which included all comic plays, were not allowed in Portugal; however, play-actors slowly introduced the *arremedilho*<sup>17</sup> in the court.

After the invention of the printing press, however, the oral tradition developed by the *arremedilho* practice declined. The Church once again invested in the popularization of the liturgical theatre. It is reasonable to infer that comic dialogues continued to be performed orally far from the castles and the churches, without leaving written registers.

---

<sup>15</sup> See Tarallo and Kato's project in which plays proved to be good sources for spoken language. Some changes are attested in plays before they are attested in other written texts (cf. Cyrino 1994, 1996; Duarte 1996; Roberts and Kato 1996).

<sup>16</sup> *Comedies of manners*: comedies that satirize social conventions (cf. Abrams 1971).

<sup>17</sup> A very simple representation based on mimes and recitation. The thirteenth and fourteenth centuries are characterized as intense periods of *arremedilhos* in Portugal.

This dichotomy between liturgical and comic plays seems to have been established in terms of themes. Thus, there is not much information on the usage of language. Notwithstanding, it is reasonable to assume that comedies staged in public places would have been performed in the colloquial language.

By the end of the fifteenth century in Portugal, two small plays written by Anrique da Mota, *Farsa do Alfaiate* and *Lamentação do Clérigo*, contained comic dialogues. After these plays, Portuguese literature was enriched by a period of high play production, known as the Vicentine period. Gil Vicente's work, mainly his comedies and farces, is distinguished by his genuine ability to translate Portugal's social reality into colloquial situations with well-elaborated characters. His dramatic plays were first performed in the court and later in public places. Frequently, however, his comedies were not allowed to be performed in the court. He is known as having founded and popularized written theatre in the sixteenth century in Portugal.

Based on the characterization of Gil Vicente's comedies and farces as reflecting colloquial situations, it was decided to investigate whether the characters used colloquial language. It was found that all the characters have dialogues, using the type of language representative of their social origins: rustic or middle class (Entwistle 1973: 351-352; Hamilton-Faria 1976: 75-96; Sletsjõe 1965: 13). The features recognized the most as being typical of oral language are related to the vocabulary and the phonology. The latter is represented in writing, by the use of such things as a high frequency of aphereses of words beginning with vowels, as in *casiam* for *ocasion* 'occasion' and *maginar* for *imaginar* 'to imagine' (Hamilton-Faria 1976: 79). Moreover, Gil Vicente regularly uses the diphthong *oi* for *ou* in order to represent the colloquial language, as in *doitor* for *doutor* 'doctor' or *poica* for *pouca* 'few' (Entwistle 1973: 353). In addition, some verbal forms are attributed to

rustic Portuguese, e.g. *trougue* ‘brought’, *prougue* ‘pleased’ and *sia* for *seja* ‘be’ (Hamilton-Faria 1976: 79).

Despite the absence of examples of the usage of non-standard syntactic variants by his characters, Hamilton-Faria (1976: 80) states that Gil Vicente was able to represent syntactic patterns other than the standard. Evidence for this is found in *Frágua d’Amor*<sup>18</sup>, in which syntactic characteristics of the Portuguese language of Negros from Guinea are revealed with impressive regularity.

Besides the linguistic features cited in literature, other features are found in Gil Vicente’s farces and comedies that are analyzed in this study, namely, double negation, as in (3); contractions, which seem to be very common in the written language of that time, as in (4); the use of diminutive forms to indicate involvement, as in (5); and the absence of prepositions, as in (6).

(3) Vistes vós!

Segundo isso, *nenhum* velho *não* tem siso natural. (Vicente 1512. VH: Moça: 71)  
‘See! / According to this, no old man has a natural sense.’

(4) Faço uma esconjuração

*c’um* dente de negra morta (...) (Vicente 1512. VH: Alcoviteira: 88) [*c’um* for *com um*]  
‘I’m making a spell / with a tooth of a dead black woman’

(5) Está tão saudosa de vós

que se perde a *coitadinha*! (Vicente 1512. VH: Alcoviteira: 99)  
‘She misses you so much/ that the poor one is sorrowful!’

(6) Esta barca *onde* vai ora, (...) (Vicente 1517. AB: Dom Henrique: 123) [*onde* for *aonde*]

‘Where does this ship go?’

---

<sup>18</sup> This work is not analyzed in this dissertation because it is labelled as a tragicomedy.

Since his work is generally recognized as containing features of the sixteenth century speech, it was decided to choose his comedies and farces as sources of data. Thus, six of his plays were selected to analyse the variation in the future temporal reference sector: *Auto da Índia* (1509), *O Velho da Horta* (1512), *Auto da Barca do Inferno* (1517), *Farsa de Inês Pereira* (1523), *Farsa dos Almocreves* (1526) and *Romagem dos Agravados* (1533).

After Gil Vicente's era, the structure and popular forms that characterized his plays were replaced by new dramatic structures. From 1580 to 1640, Portugal was ruled by Spain. During the political submission, Portuguese literature was mostly produced in Spanish or Latin, with some exceptions being written in both Spanish and Portuguese simultaneously. According to Rebello (1989: 58), a very small number of comedies were edited in Portugal during the second half of the sixteenth and seventeenth centuries<sup>19</sup>. Due to this political situation and its reflection on the written language, plays from this century are not included in this study.

Following this political submission, the Lusitane Arcadia was founded in Portugal, having as its main purpose the restoration of the national theatre. However, it replaced the Spanish influence with French classic theatre and Italian opera. The Arcadia's impact on public opinion was almost insignificant. Only the aristocracy enjoyed the glamour of the opera. During the eighteenth century, the most popular theatre included the works of the playwright António José da Silva. This Brazilian playwright lived and wrote his plays in Portugal. He was later persecuted and burned to death by the Inquisition, one reason for his death sentence was the themes of his comedies, which criticized the aristocracy's manners

---

<sup>19</sup> In the seventeenth century, the comedy *O Fidalgo Aprendiz*, written by Francisco Manuel de Melo, was considered the only important play in Portuguese theatre (cf. Rebello 1989: 59).

and conservative attitudes and contrasted them with the social condition of the Portuguese people. Two of his famous comedies, *Esopaiada ou a Vida de Esopo* (1734) and *Guerras do Alecrim e da Manjerona* (1737), are examined in this dissertation. There is no specific colloquial feature assigned to his plays in literature; however, his plays do represent colloquial situations, and they were staged for common people at the theatre of Bairro Alto (Ferraz 1980: 29-30).

In the nineteenth century, the Iberian Peninsula was invaded by France. As a consequence, Portugal's Royal Family, the court and the aristocracy fled to Brazil. From this century onwards, popular plays were written and staged in Brazil. Hence, my investigation now turns to Brazilian plays.

#### *3.3.1.1.2. Brazilian Plays*

On the other side of the ocean, the theatrical production followed the same trends as those established in the kingdom. From the beginning, Brazilian theatre was an extension of the catholic evangelization. The first author of the colony was the priest José de Anchieta (1534-1597), who produced what can be defined as sermons to be performed. He wrote in Spanish, his native language, European Portuguese and Tupi, the general language spoken on the Brazilian Coast. His plays' central theme was the celebration of the divine, especially on Christian festive dates.

According to Prado (1999), it is reasonable to suppose that from the seventeenth century onwards the play activity was intensified in Brazil as a consequence of the colonization process; however, if we examine the literary production of that century, we notice that there is no written record of play production in the colony, nor of any

performances during festivals<sup>20</sup>. Only two writers are acknowledged in the literature: Antônio Vieira, with his sermons and tracts, and Gregório de Matos Guerra, with his satirical poems. Recall that Portugal was under the domain of Spain until 1640; nevertheless, the cultural dominion continued for a far longer time, and the colony was still under religious control.

From 1760 to 1795 many theatres were built from the north-eastern coast to the extreme southern coast, namely in the states of Pernambuco, Bahia, Rio de Janeiro, São Paulo and Rio Grande do Sul. As the Italian opera spread throughout the Portuguese kingdom, it also slowly reached Brazilian theatres, and shortly afterwards, opera production was started in the colony. At first, it started merely as translation of the Italian texts imported directly from Portugal; then, it continued as plays with music performances.

As we can see, the first three centuries of theatrical work in Brazil did not make any significant contribution to the national theatre. Authors could not survive as playwrights in the colony unless they were supported by the Church or by the Portuguese court. In general, they had their formal education in Portugal and only lived part-time in the colony. Consequently, the Portuguese language they spoke and used in their written texts was not Brazilian vernacular. We cannot make substantial inferences regarding the relation between the audience and the actors in public performances, but historical fragments do mention that the play's productions were of low quality and the actors were mainly poor people of black origin. Hence, they probably did act using the vernacular, but there is no written proof of this.

In the nineteenth and twentieth centuries, the Brazilian theatre was characterized differently. The Portuguese court and aristocracy wanted the social glamour of the

---

<sup>20</sup> Twine theatre was certainly very popular in Brazil. It resembled the Portuguese *arremedilhos*.

Portuguese theatres, so they had to invest in theatrical production in the colony. They also brought companies over from Europe. A famous actor and producer of that time, João Caetano dos Santos (1808-1863), produced a large and heterogeneous repertoire as the audience was diverse. In 1838, he brought the first two Brazilian plays to the stage: *Antônio José ou O Poeta e a Inquisição* - a tragedy written by Gonçalves de Magalhães - and *O Juiz de Paz da Roça* - a comedy written by Martins Pena (Carvalho 1949: 271-273; Veríssimo 1998: 357-358).

This first comedy was used to begin my investigation on the variable under analysis. Martins Pena was the first famous comic playwright in Brazil. His comedies focused on ordinary themes about people's lives in the city of Rio de Janeiro (Veríssimo 1998: 361-362). According to Prado (1999: 56-59), this author was able to represent not only popular situations and festivities, but also people's lifestyles and ways of speaking. It is not surprising that his style was followed by two other famous comic playwrights of that century: França Júnior and Artur Azevedo. These plays were determined to be acceptable representatives of colloquial speech, following the same criterion for the selected comedies of the previous centuries.

Some examples from their comedies are listed from (7) to (14), representative of highly colloquial expressions and some linguistic features that are avoided in the standard written language. These features, among others, contributed to the characterization of their written language as representative of speech.

### **Spoken expressions/ vocabulary**

- (7) Se o doente estivesse em tuas mãos, já há muito que tinha *espichado a canela*. (Pena 1844: TM: Cautério: 271) [**for died, very colloquial**]  
 'If the sick man was under your care, he would have died long ago.'

### Usage of diminutive forms for involvement

- (8) Tua filha é um *anjinho*/ Que linda *esposinha*! (Pena 1844: TM: Lino: 255)  
 ‘Your daughter is a little angel/ What a beautiful little wife!’

### Double negation

- (9) *Não* tenho *nada*, *não*, senhor. (França Júnior 1883: DS: Catarina: 209)  
 ‘I don’t have nothing, sir.’

### Second and third possessive pronominal mix

- (10) Querem [3rd person] curar assim? Babau! Assassinos de profissão (...) desgraçados dos que se entregam em *tuas* [2nd person] mãos! (Pena 1844: TM: Aqüoso: 265) [**for suas**]  
 ‘Is that the way you want to cure it? You’re professional murderers (...) poor are those who are under your care!’

### Phonological reduction

- (11) Minha mãe preparou a jacuba *pra* meu pai? (Pena 1837: JP: Aninha: 39) [**for para**]  
 ‘Mother, have you made lunch for my father?’
- (12) Desde as quatro horas da manhã que êle saiu; está só com uma *xicra* de café. (Pena 1837: JP: Aninha: 39) [**for xícara**]  
 ‘He left at 4 a.m.; he had only a cup of coffee.’

### Non-standard pronoun placement

- (13) *Dará-se* caso que eu tenha dito alguma asneira?! (França Junior 1870: DF: Matias: 115) [**for Dár-se-a**]  
 ‘Did I say something stupid?’

### Non-standard prepositon usage

- (14) Chega *na* porta e acena com lenço. (Pena 1837: JP: author: 39) [**em for a**]  
 ‘Come into the door and wave with a handkerchief.’

In the twentieth century, the distinction between drama and comedy was often unclear in Brazilian theatre. Usually the most effective production was the one that was able to incorporate a critical analysis of society, history, families or relationships. In fact, the human condition became the central theme of provocative social plays. Historical movements, the military intervention from 1964 to 1984, and social conflicts in the country were all expressed in texts that were repeatedly forbidden.

From the twentieth century, five plays were selected from four Brazilian playwrights: Chico Buarque, Gianfrancesco Guarniere, Dias Gomes and Aimar Labaki. As in the previous plays, those of the twentieth century included dialogues about daily events. The language used by the above authors in their plays consisted of many linguistic features which represented the way in which people actually spoke and thus did not necessarily conform to the standardized forms of grammarians. From (15) to (23), are passages from these plays that justify their selection for this study.

### Spoken expressions/ vocabulary

- (15) Matilde, tu sabe como é esse padre. Vai fazer um *fuzuê*. (Gomes 1963: BH: Rosali: second frame: 393)  
 ‘Matilde, you know what this priest is like. He is going to make a commotion.’
- (16) Vamos *molhar um pouco a goela* na venda do seu Demerval, Zelão. (Gomes 1963: BA: Ambrósio: first frame: 222)  
 ‘Zelão, let’s have something to drink at Demerval’s store.’

### Phonological reduction

- (17) Mãe, o pai *tá* comprando na barraca daquela perua! (Gomes 1963: BH: Mocinha: second frame: 384) [for *está*]  
 ‘Mother, my father is buying things from that bitch’s stand!’
- (18) *Inda* não vendeu tudo? (Gomes 1963: BH: Malta: second frame: 389) [for *ainda*]  
 ‘Haven’t you sold everything yet?’

- (19) Ou você consegue um lugar, *num* desses prédios da Corporação, *pra* essa gente se abrigar ou eles vão morrer de frio, de fome. (Labaki 1999: B: Verônica: scene 4: 55) [**for *em um* and *para*, respectively**]  
 ‘Either you find a place for sheltering them in one of the Corporation building, or they will die of cold and starvation.’

### Second person verbal non-agreement

- (20) Olha, Tião, são dois anos que *tu* não *acerta* as contas comigo. (Buarque 1978: OM: Max: scene 2: 57) [**for *tu acertas***]  
 ‘Look Tião, you haven’t payed your debts for two years now.’

### Non-standard usage of third person subject pronoun

- (21) Olha, lá, tá sangrando! Vamos tirar *ele* de lá. Vamos tirar. (Guarnieri 1961: S: Américo: second act: scene 4: 76) [**for *Vamos tirá-lo de lá.***]  
 ‘Look over there! He is bleeding! Let’s take him out of there. Let’s take him.’

### Non-standard pronoun placement

- (22) *Me* deixa, Cipriano! E fica sabendo, não vou a passeata nenhuma! (Guarnieri 1961: S: Américo: second act: scene 3: 71) [**for *Deixe-me***]  
 ‘Leave me alone, Cipriano! I won’t go to any protest, and that’s it!’

### Non-standard preposition usage

- (23) Quando chegar *em* Salvador, vou fazer uma propaganda danada disso aqui. (Gomes 1963: BA: Ernesto: fifth frame: 276) [***em* for *a***]  
 ‘When I arrive in Salvador, I will promote this city.’

As already stated, Brazilian comedies were only staged from the nineteenth century onwards. Table 3.1 displays the selected plays, the centuries, their codes, the years, their authors, and the provenance of the plays and the authors. PRT stands for Portuguese and BR stands for Brazilian (cf. Appendix 2 for a complete bibliographic reference).

*Table 3.1: Selected Plays*

PLAYS						
Century	Code	Year	Title	Author	Provenance Play	Author
XVI	AI	1509	Auto da Índia	Gil Vicente	PRT	
	VH	1512	O Velho da Horta	Gil Vicente	PRT	PRT
	AB	1517	Auto da Barca do Inferno	Gil Vicente	PRT	
	IP	1523	Farsa de Inês Pereira	Gil Vicente	PRT	
	FA	1526	Farsa dos Almoceves	Gil Vicente	PRT	
	RA	1533	Romagem dos Agravados	Gil Vicente	PRT	
XVIII	VE	1734	Esopaiada ou A Vida de Esopo	Antonio José da Silva (o Judeu)	PRT	BR
	AM	1737	Guerras do Alecrim e da Mangerona	Antonio José da Silva (o Judeu)	PRT	
XIX	JP	1837	O Juiz de Paz da Roça	Martins Pena	BR	BR
	D	1842	Os Dous ou O Inglês Maquinista	Martins Pena	BR	
	TM	1844	Os Três Médicos	Martins Pena	BR	
	N	1853	O Noviço	Martins Pena	BR	
	DF	1870	O Defeito de Família	França Júnior	BR	BR
	TB	1872	O Tipo Brasileiro	França Júnior	BR	
	DS	1883	Dois Proveitos em um Saco	França Júnior	BR	
	LB	1885	A Lotação dos Bondes	França Júnior	BR	
	CF	1897	A Capital Federal	Artur Azevedo	BR	BR
XX	S	1961	A Semente	G. Guarniere	BR	BR
	BA	1962	O Bem Amado	Dias Gomes	BR	BR
	BH	1963	O Berço do Herói	Dias Gomes	BR	
	OM	1978	Ópera do Malandro	Chico Buarque	BR	BR
	B	1999	A Boa	Aimar Labaki	BR	BR

Having considered Labov's comments on historical and literary language sources, it was important to find written language that showed some regularity and oral features. How did the written Portuguese language differ from the spoken variety in these centuries, especially in terms of the variable rule examined in this dissertation? Unfortunately, there is

no guarantee that plays from either one or five centuries ago could really be representative of the spoken Portuguese language, nor can we know whether the language spoken on the stage was the same as the variety used in the written version. Even so, the investigation into the literary production in Portugal and Brazil, into the production of plays, into the way in which the socio-political events affected this production, as well as into the various styles of the authors, and, most importantly, the linguistic features of these plays support their selection as representative of speech, though it is necessary to remember, of course, that the natural, uncontrolled language is not obtained from written sources.

### *3.3.2. The Synchronic Data*

#### *3.3.2.1. Brasilia*

The synchronic data were collected in the city of Brasilia, Brazil. This city, the capital of the country, is located in the west-central region of Brazil and was built in the sixties to be an administrative medium-sized city.

During the sixties and seventies, there was an increase in the demographic density of the city, which was formed by immigrants who came mainly from the north-eastern region of the country as well as from the states of Rio de Janeiro and Minas Gerais, which are located in the south-east. Among these areas, the dialect of Rio de Janeiro is considered as one of the most prestigious in the country due to the influence of the city being the capital from 1763 to 1960. The first administrative employees of Brasilia were transferred from Rio de Janeiro, while the builders came from the north-eastern region. The majority of working-class and lower middle-class speakers live in the so-called satellite-towns, i.e. suburbs. After four decades, however, the city has reached 2 051 146 inhabitants. It presents an unequal income distribution and recently some slums were formed, mainly on its borders

with the state of Goiás. Its slums are not as precarious as the ones in the rest of the country; however, they share the same problems, namely, a high degree of illiteracy and unemployment.

In attempting to characterize some of Brasília's linguistic features, the most salient is its unmarked pronunciation. As Bortoni-Ricardo (1991) states, a merging of regional and social linguistic features may have led to a restructuring of the speakers' phonological system in such a way that they replace some of their salient regional variants by unmarked ones. In the social structure of the city, there is neither strict rigidity, nor a well-established ideology of ethnic group loyalty. Consequently, the immigrants' descendents do not have strong feelings about maintaining the dialect. Brasília's accent is not marked by any of the well-defined regional accent features of Brazilian Portuguese.

However, Brazilian regional accents are evaluated differently in the city: the south-eastern and southern accents have more prestige than the north-eastern accent; even though, the great amount of immigrants living in the city came from the latter region. The determining factor in this evaluation is probably the economic development of the regions: the accents from the poorest regions are associated with lower social status.

As there is a natural inclination to recognize the capital as the political and cultural centre of the country, the educational system and the media tend to focus on the local, unmarked dialect. This linguistic aspect of Brasília confers relative awareness of phonological marked features by speakers. Nevertheless, this characteristic of the city does not affect the variable under analysis in this study.

### 3.3.2.2. *The Interviews*

The first step of a variationist study is to have available a reliable source for the gathering of the data. In the case of the synchronic oral data of this dissertation, this source consists of tape-recorded interviews with working class Brazilian Portuguese speakers, and more specifically, with urban speakers from the country's capital. All informants were drawn from elementary, junior and upper school levels.

The sample is formed by a group of informants some of whom knew each other from school. With the exception of five informants<sup>21</sup>, all the interviews were conducted personally by the author. The informants were contacted in the neighbourhood or at the university, either by the author or by one of her former students. Each informant was interviewed once either in their home or in another place of their choosing. So as to make the interview as informal as possible, at least one person who knew the interviewee was present during the tape-recording.

The interviews were at least sixty minutes in length, though a few were longer. All interviews followed the same list of conversational topics, being, however, flexible for new topics that could be proposed by the informants. The interview contained modules about school, teachers, and topics related to the educational system. Then, they were encouraged to express their thoughts about the social condition of the city, their daily lives, their neighbourhood, and their personal experiences of holidays, festive dates, danger of death, religion, mystical experiences, dreams, memorable movies, family and childhood. Some informants were naturally more talkative, so there was almost no need to introduce new

---

<sup>21</sup> Five informants from the sample were not interviewed by me, namely, Nataly, Elaine, Neila, Eládio and Natália. However, the interviews followed the same conversational topics.

topics or give feedback; others had to be guided to some specific topics, but generally became more cooperative as the recordings were being done.

These topics were selected with the intention of making conversation as natural as possible. The aim was to make the informants be involved in the content of their speech so that they would not pay attention to or control the linguistic forms and structures they were using. This is a way of establishing a rapport between the interviewer and the informants and minimizing the Observer's Paradox: "to observe the way people use language when they are not being observed" (Labov 1972: 61; Guy 1996: 242-45). Only by means of minimizing this paradox can the sociolinguistic researcher access the vernacular, i.e. language as it is really used by speakers in their daily lives.

#### *3.3.2.3. The Sample*

The group is basically constituted of young and young adult informants who were born and/or raised in the capital. It cannot be considered statistically representative of the population of the whole Federal District because informants from rural areas, upper social class, illiterate and elders were not incorporated in the sample. However, the informants represent the social classes that hold the majority of the population of the city, and in a larger scale the majority of Brazilians, i.e. the young and young adults in the working and lower middle social economic classes.

The informant selection also followed an educational background criterion. Three groups were formed with the objective of having informants whose language had been exposed to the influence of formal education. The first group was from primary level, which comprised of the first six grades of schooling; the second one was from an intermediate level, which corresponds to high school; and the last group was from

university. Thus, all informants were under the influence of schooling, which is the traditional vehicle for disseminating the standard language in the country. With respect to the variable under analysis in this study, the prescriptive standard is the usage of the synthetic future, which is selected for conjugation table, and which is acknowledged the most as meaning futurity in the language (cf. chapter 2).

The subdivision of the sample into three educational levels was designed to answer one question in this research: does the selection of the synthetic future as the standard affect the usage of this variable in colloquial language? If the grammatical tradition has successfully echoed in Brazilian Portuguese, the educational-level may be reflected in the speaker's usage of the synthetic future. Thus, the more speakers receive formal education, the higher the probability is of choosing the synthetic future in their natural speech. On the other hand, if the grammatical tradition has not been able to influence natural language usages, there will be no correlation between speakers' educational level and the choice of the synthetic future variant.

Table 3.2 shows the sample constitution of the corpus of spoken Brazilian Portuguese:

*Table 3.2: Informants*

Nº	Informant	Age	Gender	Schooling	Year of Interview	Years in Brasilia
01	Ana	10	F	4 <sup>th</sup> grade	1991	10
02	Ana Paula	10	F	4 <sup>th</sup> grade	1997	10
03	Karina	10	F	4 <sup>th</sup> grade	1991	10
04	Lilian	10	F	4 <sup>th</sup> grade	1991	10
05	Nina	10	F	4 <sup>th</sup> grade	1991	10
06	Priscila	10	F	4 <sup>th</sup> grade	1991	10
07	Diego	10	M	4 <sup>th</sup> grade	1992	10
08	Edigard	12	M	4 <sup>th</sup> grade	1992	12
09	Jose Luiz	13	M	4 <sup>th</sup> grade	1991	13
10	Tiago	10	M	4 <sup>th</sup> grade	1991	10
11	Víctor	10	M	4 <sup>th</sup> grade	1997	10
12	Filipe	12	M	6 <sup>th</sup> grade	1997	12
13	Érica	14	F	8 <sup>th</sup> grade	1992	14
14	Gabriela	14	F	8 <sup>th</sup> grade	1997	14
15	Jeanne	14	F	8 <sup>th</sup> grade	1992	14
16	Karina	17	F	8 <sup>th</sup> grade	1992	17
17	Vânia	14	F	8 <sup>th</sup> grade	1992	14
18	Cássio	14	M	8 <sup>th</sup> grade	1992	14
19	Salvador	15	M	8 <sup>th</sup> grade	1992	15
20	Suziel	14	M	8 <sup>th</sup> grade	1992	14
21	Patrícia	21	F	10 <sup>th</sup> grade	1997	21
22	Nataly	27	F	11 <sup>th</sup> grade	2003	27
23	Neila	28	F	11 <sup>th</sup> grade	2003	28
24	Alessandra	20	F	University	1992	17
25	Cátia	19	F	University	1992	19
26	Elaine	28	F	University	2003	28
27	Eliane	18	F	University	1992	18
28	Luciana	25	F	University	2003	24
29	Natália	30	F	University	2003	30
30	Cláudio	20	M	University	1992	18
31	Eduardo	23	M	University	1992	20
32	Eládio	19	M	University	2003	19
33	Rodrigo	25	M	University	2003	24
34	Teotônio	21	M	University	1992	21

## Chapter 4 *Analysis*

---

The last two chapters laid the groundwork for this research by (i) describing the representation of the future tense in language, including the meanings that are ascribed to the future forms by prescriptive works; (ii) outlining the theory and the methodology that frame this work; and (iii) introducing the data set submitted to the variable rule analyses. The next step in this study is to define the variable context in detail by characterizing the expressions of futurity in the language, as well as the interchangeable and non-interchangeable contexts where these occur, the working hypotheses about their use, and the discrete factors operationalized to test them against the data.

### **4.1. The Linguistic Variable**

As presented in chapter 2, the future tense in Portuguese can be expressed through multiple forms, which are generally considered to convey different pragmatic and semantic meanings other than strictly temporal reference. The forms are considered to be related to each other through a common abstract concept, i.e. futurity, but with distinctions.

The nuances of meaning associated with these different forms can be tested empirically by quantifying the variable use of the forms in contexts where they are interchangeable (Poplack 1992; Poplack & Turpin 1999; LeBlanc & Poplack 1999; Poplack & Tagliamonte 2000; Poplack & Tagliamonte 2001). Whenever possible, the hypothesized meanings are operationalized as factors that can be quantified in the data.

It is critical in such an investigation to distinguish the formal uses of future tense forms from actual functional reference to future events. In addition to temporal reference some future forms can be used in other contexts, as for example the use of the synthetic future to express the imperative (cf. section 2.3.2.1.). A review of the prescriptive literature and a close inspection of the data revealed many meanings which were not compatible with a temporal reading as observed in section 2.3.4.

The study of a variable rule requires that the forms are used in contexts interchangeably; so that, if there are differences among them, in the variable context the differences are not applied, i.e. they are neutralized. Since the central goal of this dissertation is to ascertain the constraints that determine the usage of future forms in order to perceive how the future temporal sector has developed in the Portuguese language, only interchangeable forms in future temporal reference contexts are analyzed. Hence, this study takes as its locus of variation the future temporal reference context, and has as its departure point the context itself, instead of the forms.

The future temporal reference in the Portuguese language is expressed in the corpus studied by four variants, which are exemplified in (1) to (4):

**(1) Synthetic Future (SF)**

Ai, que cousa me *pedirá* (SF) vossa mercê que lhe não faça, dizendo-me onde estão as minhas meninas. (Silva 1737. AM: Fagundes: scene 1: 46)<sup>22</sup>

‘Whatever you *ask* of me, I will do; if only you would tell me where my daughters are.’

---

<sup>22</sup> The information in parenthesis is respectively the author and year of the play, the play code, the character, the part and the page where the token is located. If it is a spoken datum, the page may be exchanged by the counter number.

(2) **Periphrastic Future: *haver* (present form) + *de* + infinitive (HP)**

E fie de mim que o meu prêmio *há-de ser* (HP) igual ao meu desejo. (Silva 1737.

AM: D. Fuas: scene 1: 47)

‘Trust me. My prize *will be* equal to my desire.’

(3) **Present (P)**

Eu tou te ligando pra lembrar que amanhã é o último dia do mês (...) sua dívida tá em 30 contos e no primeiro dia do mês *passa* (P) a 33. (Buarque 1978. OM. Duran:

scene 1: 23)

‘I’m calling to remind you that tomorrow is the first day of the month (...) your debt is 30 bucks and *will increase* to 33 on the first day of the month.’

(4) **Periphrastic Future: *ir* (present form) + infinitive<sup>23</sup> (GP)**

Quantas filhas que criei / eu as pus em pobre sorte / *Vou morrer* (GP),/ Elas hão de padecer, / porque não lhes deixo nada; / de quanta riqueza e haver / fui sem razão despende / (Vicente 1512. VH: Velho: 105)

‘Many daughters I nurtured / I left them in a hard luck / *I’m going to die*, /They will suffer, / because I don’t leave anything to them; / how much wealth and possessions / I spent without a reason/’

All four of these variants can be used interchangeably in future temporal reference contexts, however not all are equally well represented in all stages and varieties of Portuguese. For example, the variant in (2) is not considered a future temporal marker in

<sup>23</sup> In Portuguese, there is also the usage of the synthetic future form of *ir* ‘to go’ + infinitive in the future temporal reference context; however, this usage was virtually not attested in the data in all centuries (only seven occurrences). It is exemplified below:

**Periphrastic Future: *ir* (inflected future form) + infinitive**

Se não, eu ficarei louco e *irei viver* na Hircânia, no Japão, em toda parte aonde possa encontrar desumanas feras, e lá morrerei. (Pena 1842. Cecília: scene 9: 44)

‘If not, I will be crazy and *will live* in Hircânia, in Japan, wherever I can find beasts, and there I will die.’

contemporary Brazilian Portuguese. A detailed distribution of variants by century will be given in chapter 5.

#### 4.2. Circumscribing the Variable Context

The identification of the context for analysis is one of the crucial stages of the variationist study. It is critical that the context include only true variants, i.e. variants that can be used interchangeably without any appreciable difference in meaning. Furthermore, only studies where the variable context has been rigorously and consistently defined can be assured of comparable results across studies and corpora. Inclusion of tokens which are not qualified as variants will interfere in the statistical analysis and hamper comparability of results.

Since the locus of this study is future temporal reference in Portuguese, all non-future contexts must be excluded from the quantitative analysis, as well as any invariant occurrence of a future form. A number of potential variants were excluded from the variable context:

The non-future uses include such forms as imperatives (5) (including directives found in recipes, instructions, formulas etc.); hypothetical questions (6); and, in (7) lexical uses of *ir* + infinitive when the context clearly indicates movement toward another place<sup>24</sup>.

---

<sup>24</sup> Thomas (1969, 1974) claims that the time indicated in these cases is always future, even in the sentences in which an implication of movement is implied. In his own words: "Use of the form *vou falar*. This is very frequent use in BF, as in English and several Romance languages. While it may have the meaning of movement toward another place in order to perform an action, in most cases this original meaning of the verb *ir* is completely gone, so that the phrase is essentially a future form of the verb represented by the infinitive. In many cases it is impossible to determine whether the speaker has in mind an idea of movement or only futurity. However, even when movement is clearly implied, the time indicated is still future." As stated in this section, however, I excluded all *ir* + infinitive forms when a sense of movement was implied by the context.

- (5) Sim, Esopo, tu *hás de ensinar* a esta rapariga a ler, e a escrever; aí t'a entrego. (Silva 1734. VE: Xanto: first part, scene 2: 37)  
'Yes, Esopo, *teach* her to read and write; then I will give her to you.'
- (6) Tá muito quieto mesmo. *Será* que ele está bom? (Guarnieri 1961. S: João: act 1: scene 3: 37)  
'It's really too quiet. *Would* he be fine?'
- (7) [Aonde vais?]  
*Vou mudar* de camisa, que entendo, que estou mijado com alguma coisa mais. (Silva 1734. VE: Esopo: second part, scene 5: 122)  
'[Where do you go?] I'm going to *change* my shirt, because I think, I'm wet with something else.'

Olha cá, Fernando, eu *vou/ ver* a com que hei de casar, / avisa-te que *hás de estar* / sem barrete onde eu estou. (Vicente 1523. IP: Escudeiro: second part: 202)  
'See, Fernando, I'm going to *see* her, to whom I will marry, / tell her that she must be/ where I am.'

Other non-future usages include *ir* in the present tense corresponding to the English expression 'Let's', as in (8) below,

- (8) Hoje a gente vai relaxar. *Vamos ler* um pouquinho pra gente relaxar, né. Porque dois horários de Português é da pesada. (1992. Vânia: A: 21)  
'Today we're going to relax. *Let's read* a little bit just to relax because two classes of Portuguese are too much.'

And hypothetical contexts, as the protasis of the conditional *se* sentences which were also excluded, as in (9) below:

- (9) Não sou nenhum canalha. *Se minha presença aqui vai lhe causar algum problema*, faz de conta que a gente não se conhece. (Gomes 1963. BH: Roque: act 1: frame 6: 425)  
'I'm not a scoundrel. *If my presence is going to be a problem*, we behave as not knowing each other.'

Besides the contexts cited above, the future forms were also found in habitual contexts. These forms denote habitual actions or general truths. They were not considered as occurring in a future context and, consequently were excluded.

- (10) Você lê, todo mundo lê o livro, aí depois *vamos debater*, aí debate o livro, isso é uma aula boa porque tá debatendo. (1992. Salvador: A: 25)  
 ‘You read, everybody reads the book; then we *debate* it, so we debate the book. That is a good class because we debate.’

Also certain numbers of set expressions like *vai ver* and *you contar*, as in (11) and (12), were excluded.

- (11) Beleza, você *vai ver* se não tomar cuidado. (Guarnieri 1961. S: João: act 1: scene 3: 33)  
 ‘Girl, if you do not take care you *are going to see*.’
- (12) Você deve tá com solitária, compadre. Você come que eu *vou te contar*! (Guarnieri 1961. S: Operário 1: act 2: scene 3: 67)  
 ‘You must have a tapeworm, buddy. *Let me tell* you how much you eat!’

As usual in variationist studies, proverbs, citations, quotations, metalinguistic commentaries and self-corrections were not considered as productive, and were not included in the quantitative analysis.

- (13) Diz o exemplo da velha: / “o que não *haveis de comer* / deixai-o a outrem mexer”. (Vicente 1523. IP: Inês: first part: 199)  
 ‘The old woman saying says: / what you *will* not eat / leave it to others.’
- (14) Porque de pedra *será* minha igreja e você *será* seu fundamento. (Labaki 1999. B: Ricardo: scene 1: 13)  
 ‘Because my church *will be* of stone and you *will be* its foundation.’

- (15) Eu acertei qual *vai ser* a espe/ qual *será* a espessura de um esgoto. Aí eu, cara, eu não consigo me ver fazendo isso. (1992. Eliane: A: 5)  
 'I've got right what *is going to be* the thick/ what *will be* the thickness of the sewer. Boy, I can't see myself doing this.'

#### 4.3. Previous Quantitative Works on the Future Forms in Brazilian Portuguese

Two previous studies on the use of future forms have been carried out on Brazilian Portuguese. The first one was a quantitative, but not variationist work on the use of future forms by university students from the state of São Paulo (Baleeiro 1988); the second study was a variationist analysis of the future temporal reference contexts in formal and written language of Parliament discourses and magazine articles (Santos 1997).

Baleeiro analyzed each occurrence of a future form that emerged from twenty-four hour recording of formal dialogues. Regarding the usage of future forms in future temporal contexts, she attests a total of 77% of *ir* + infinitive periphrasis usage, in contrast to 13% of synthetic forms, from a total number of 453 occurrences. The present form and the *irei* + infinitive periphrasis were also found in the corpus, 9% and 1% of occurrences, respectively.,as shown in Table 4.1:

**Table 4.1: Distribution of future variants in formal spoken Brazilian Portuguese.  
 (adapted from Baleeiro 1988)**

Variants	%	Number of tokens
Synthetic future	13	59
<i>Ir</i> + infinitive periphrasis	77	349
<i>Irei</i> + infinitive periphrasis	9	40
Present	1	5
Total		453

Although Baleeiro's study does not attempt to analyze the conditioning factors that favour one or another variants, she does make a number of observations about the data that can be operationalized as factors for this study. First, she observed the tendency of the synthetic future and the present form to co-occur with other elements expressing future time, e.g. time adverbs, as opposed to the periphrastic form *ir* + infinitive that was generally used without such elements in the sentence (cf. Baleeiro 1988: 73, 110 and 180). Second, she observed the association of the third verbal person and [-human] feature with the use of the synthetic form (op. cit.: 177).

Santos' (1997) work on the future temporal reference is based on a formal corpus of written and spoken language, mostly acquired from political discourse and articles from a well-known national magazine. The distribution of the future forms in her research is as follows:

***Table 4.2: Distribution of future variants in formal written Brazilian Portuguese.  
(adapted from Santos 1997: 45)***

<b>Variants</b>	<b>%</b>	<b>Number of tokens</b>
<b>Synthetic future</b>	<b>73</b>	<b>772</b>
<i>Ir</i> + infinitive periphrasis	22	233
<i>Irei</i> + infinitive periphrasis	3	31
Present	2	19
Total		1055

When comparing Santos' distribution of forms to that of Baleeiro, it is interesting to notice that there is inversion in the ranking of the two mainly used variants. The synthetic future accounts for 73% of the tokens in Santos' work (cf. Table 4.2.), while *ir* + infinitive accounts for 77% of the tokens in Baleeiro's (cf. Table 4.1.). As both corpora are said to be formal, I assume that the difference in their distribution is attributed to the written language; especially because even the spoken data in Santos' research was transcribed into a written form by employees of the Parliament or by journalists.

Santos' variable rule analysis of the synthetic form and *ir* + infinitive variants (N=991) found the following factors to favour the use of the synthetic future variant, (1) special verbs, as *to be* and *to have*; (2) previous occurrence of the synthetic form; (3) third grammatical person; (4) monosyllabic verbs, and (5) the most formal data, i.e. data gathered from the Parliament as opposed to magazines. Although Santos tested the influence of the use of temporal adverbs in the selection of variants, this factor was not selected by the statistical analysis.

Baleeiro and Santos identified in their studies factors that had been remarked upon by prescriptive grammarians, such as *number of syllables*, *adverbial specification*, and *grammatical person*. These factors will also be tested on the spoken data in the current study to allow for a comparison with their findings from formal registers.

#### 4.4. Linguistic Hypotheses, Factor Groups and Coding

Investigation of variable phenomenon is essentially data driven, in that the analyst generally elaborates some basic hypotheses on the behaviour of the variants. This study, however, employs an additional methodology elaborated by Poplack et al. (2001). This approach traces references to a particular form through the history of written grammars and extracts their comments as a source for investigation, as detailed in chapter 2.

Poplack et al. (2001: 93) advocates a three-pronged approach in order to obtain valuable results for variationist studies when investigating the variant forms mentioned in grammar books, namely, “(1) evaluation of the unequal treatment of forms in grammars, (2) interpretation of evidence of the variability, and (3) tracing evidence of the conditioning of that variability”. This study concerns itself primarily with the third point. All comments on the meanings and the usage of the future forms in the future temporal reference context that could be tested empirically were operationalized as hypotheses and, accordingly, as factors for the variable rule analyses. Moreover, previous variationist works on the topic and remarks about the future in Portuguese when provided by non-prescriptive literature were also considered at this stage of the work.

Each token was coded according to working hypotheses of this research. Each hypothesis corresponds to one factor group, which is comprised of independent factors. These factor groups are listed in (16).

(16) *Factor Groups Coded**Lexical-Semantic and Morphosyntactic:*

Number of syllables of main verb stems (one, two, three or more)

Stativity of main verb (stative, motion)

Grammatical person and Number of subject

(first singular, first plural, second singular, second plural, third singular, third plural, inexistent subject)

Animacy of subject (human, animate, inanimate, inexistent subject)

Presence of clitic pronouns

(absence, direct object, indirect object, both, reflexive and passive)

*Syntactic and Discourse:*

Polarity (negative, affirmative)

Sentence type (declarative, question)

Presence and Type of adverbial specification

(no adverbial, specific, non-specific)

Type of clause (main clause, subordinate clause, relative clause)

Contingency of event (contingent, assumed)

Temporal distance (distal, proximal)

*Extralinguistic factors:*

Century: (1500, 1700, 1800, 1900)

Medium (written, spoken, 20<sup>th</sup> C only)

Level of education in the spoken data (primary, high school, university)

These factors are related to the following hypotheses from the prescriptive and sociolinguistic literature.

1 – *Number of syllables of main verb stems*: One-syllable main verbs have been considered the locus of the realization of the synthetic future in the Portuguese language. It has been claimed that the synthetic form usage is mainly restricted to monosyllabic verbal stems (Thomas 1969, cf. section 2.3.3.1.). In order to test this hypothesis, each token was coded according to the number of syllables of the main verb: one syllable verbs, as in (17); two syllable verbs, as in (18); and three or more syllables, as in (19):

- (17) Filena, pois sempre amou Periandro, casem que eu *serei* (SF) o padrinho, (Silva: 1734. VE: Esopo: second part: scene 11: 155)  
‘Filena, you always loved Periandro, get married that I *will be* your best man,’
- (18) Senhor, eu o *servirei* (SF) e não quero outro senhor, (Vicente: 1526. FA: Ouríves: 183)  
‘Sir, I *will serve* you and I don’t want other master,’
- (19) Não me *levantarei* (SF) enquanto mo não deres. (Pena 1853. N: Ambrósio: second act: scene 5: 98)  
‘I *won’t stand up* while you don’t give it to me.’

2 – *Stativity*: The periphrasis *ir* + infinitive has been claimed to signify a movement to perform an action in a future time (Bechara 1968; Said Ali 1969; Thomas 1969; Cunha & Cintra 1984/ 1999, cf. section 2.3.). According to Poplack and Tagliamonte (2000: 335), the construction of this periphrasis with verbs of motion entails the ‘bleaching or desemanticization of its original lexical content’. If indeed *ir* retains some nuance of its original lexical meaning, then one would expect it to be less likely to occur with other verbs of motion. Verbs were coded as being of movement, as in (20), or not, as in (21):

- (20) Eu não tou com medo, e vou andando. Eu não *vou correr* (GP), senão não me chamo Vânia. Eu vou andando, sabe? (1992: Vânia: B: 44)  
‘I’m not scared, and I’m walking. I *won’t run*, or I’m not Vânia; and I’m walking, you know?’

- (21) Só que tem muita gente que *vai reprovar* (GP), por causa de bagunça e de nota. (1992. Priscila: A: 14)  
 ‘There are many students who *will fail* because of their behavior and their grades.’

3 – *Grammatical person and number of subject*<sup>25</sup>: According to some grammarians, both the periphrastic forms and the present convey certainty, conviction, and volitional meanings (Vieyra 1809; Soares Barboza 1852, in section 2.3.3.2; Cunha & Cintra 1984/1999; Arruda 2000; Kury 1986; Sardinha & Ramos 2000; Thomas 1969, 1974; Rossi 1945, cf. sections 2.3.3.3. and 2.3.3.4.). Conversely, certainty has also been claimed as one of the nuances of meaning expressed by the synthetic form (Thomas 1969; Cunha & Cintra 1984/1999, cf. section 2.3.3.1.). If the usage of the future forms is related to speakers’ intent in the Portuguese language, it is expected to find a relation between first person and the choice of future forms, as certainty is most strongly asserted by the speaker. The choice of periphrastic forms and the present usage with first persons in the informal data would be consistent with the findings of Baleeiro (1988) and Santos (1997) for formal Brazilian Portuguese. This hypothesis was tested by coding the tokens according to number (singular, plural) and the grammatical person (first, second, third) as well as for inexistent subjects, as in (25). Later, the number distinction was collapsed; only the grammatical person distinction was maintained.

- (22) Quando eu for em Petrópolis, eu *vou gostar* (GP). (1992. Cláudio: A: 02)  
 ‘When I go to Petrópolis, I *will like* it.’

- (23) Tu me *pagarás* (SF)! (Pena 1842. JP: scene 1: frame 2: 22)  
 ‘You *will pay* me!’

---

<sup>25</sup> The factor groups of *grammatical person* and *animacy* showed a high degree of interaction. They will be collapsed for the statistical analyses. See section 4.5 for details.

- (24) O presidente lançou um comunicado lá dizendo que não *vai reclassificar* (GP) ninguém. (1992: Alessandra: A: 04)  
 ‘The president sent a memorandum saying that (he) *will not promote* anybody.’

- (25) *Vai fazer* (GP) um ano. (1992. Eliane: A: 1)  
 ‘(It) will complete one year.’

4 – *Animacy*: This is another way to try to objectively operationalize the notion of speaker certainty and volition. Volition, intention, certainty, and desire are all related to the agentivity of the subject. The [+human, +animate] subject specification confers a higher degree of certainty and compromise to the verbal action. It is therefore expected that periphrastic and present forms will be favoured with [+human, +animate] subjects; while [- human], inanimate and inexistent subjects would favour the synthetic form. In order to test this hypothesis, the subject was coded as human, as in (26); animate, as in (27); inanimate, as in (28); or inexistent subjects, as in (29):

- (26) Mas já que sua filha casou mesmo, *você não vai dar* (GP) um presentinho? (Buarque 1978. OM: Geni: scene 1: 39)

‘But once you daughter already got married, aren’t you *going to give* her a little gift?’

- (27) Se o vento tá batendo pra você. Tá vindo assim na sua direção. *Aí, ele* [o animal] não *vai sentir* (GP) meu cheiro. (1997. Gabriela: 2A: 242)

‘If the wind is coming from your side, it is coming toward you, then, it [the animal] *is not going to smell* me.’

- (28) Se eu fizer isso, *a notícia vai se espalhar* (GP) e em menos de uma semana vai ter uma multidão querendo entrar aqui. (Labaki 1999. B: Ricardo: scene 4: 58)

‘If I do this, *the news is going to spread out*, and there is going to be a crowd trying to get in here in less than one week.’

- (29) *Aí eu, que festa?* Ah, na nossa sala *vai ter* (GP) uma festa. *Aí, eu, ah é?* (1992. Tiago: A: 2)

‘So I, well, which party? Well, there *is going to be* a party in our classroom. Then, I, really?’

5 – *Presence of clitic pronouns*: The placement of clitic pronouns with the synthetic future form has been considered problematic in the Portuguese language (cf. Nunes 1910; William 1968; Melo 1938, in section 2.2.; also Barreto 1944; Machado Filho 1938; Figueiredo 1941, in section 2.3.3.2). The prescriptive orientation for their placement puts them in the middle of the synthetic future form, e.g. *dar-lhe-ei*, ‘I will give [it] to you’. This creates a mesoclesis, which has been considered as artificial in the language (cf. Nunes 1910, section 2.2.), but is still prescribed. It is plausible therefore, that speakers would avoid the usage of synthetic forms in contexts with clitics. Thus, I expect that periphrastic and present forms will be favoured in contexts with clitics in Portuguese. The effect of clitics on future tense selection has never been quantitatively studied before in Portuguese, so it was determined to make a number of fine-grained distinctions. Tokens were coded as occurring with no clitics, as in (30); with direct object clitics, as in (31); with indirect object clitics, as in (32); with both direct and indirect object clitics, as in (33), and reflexive or passive pronouns, as in (34):

- (30) Eu que falo com ela, e muito, sei que ela não *há de aceitar* (HP) proposta de casamento sem estar muito bem coberta. (Buarque 1978. OM: Vitória: scene 1: 32)  
‘I talk to her a lot. I know she *won't accept* a marriage proposal without guarantees.’
- (31) Pois aonde *o hei de pôr* (HP)? (Silva 1737. VE: Semicúpio: scene 2: 54)  
‘Where *will I put* it?’
- (32) Ela *lhe dará* (SF) o recado. (Silva 1734. VE: Periandro: first part: scene 2: 34)  
‘She *will give you* the message.’
- (33) Não *vo-lo hei de negar* (HP): fazei-me uma petição. (Vicente 1526. FA: Fidalgo: 46: 176)  
‘I *won't deny it to you*: make your request.’

- (34) Ora, faça-me um favor! Então esse pessoal que já enfrentou metralhadora *vai se assustar* (GP) por causa de cinco tiras! (Guarnieri 1961. S: Agileu: first act: scene 1: 28)  
 ‘Oh, come on! These people who faced machine-guns *aren’t going to be scared* of five cops!’

At the syntactic and discourse level, this study investigates the effects of: sentence polarity, sentence type, adverbial specification, type of clause, contingency, and temporal distance, as listed in (16). The hypotheses are formulated as follows:

6 – *Sentence polarity*: Soares Barboza (1852, cf. section 2.3.2.2.) commented that the synthetic future form was used in the Portuguese language to express a positive future time. If Soares Barboza’s remark is correct, then affirmative statements should show a tendency to favour the synthetic future in Brazilian Portuguese. This hypothesis was also elaborated due to the selection of this factor group as statistically significant in the choice of the future variants in French (Poplack & Turpin 1999); however, in that case negation strongly favoured the synthetic future occurrence. To test the influence of *polarity*, sentences were coded as affirmative or negative. The examples in (35) and in (36) shows negative and affirmative sentences, respectively:

- (35) Ah não, a gente **num** *vai conseguir* (GP) ganhar, né; então, *vou deixar* logo de lado.  
 (1992. Gabriela: A: 17)  
 ‘We’re **not going to win**; so, I’m *going to give up*.’
- (36) Senhor, eu o *servirei* (SF) e não quero outro senhor, (Vicente 1526. FA: Ourives: 202: 183)  
 ‘Master, I will serve you and I don’t want another master,’

7 – *Sentence type*: Sentence type has been considered as relevant to the selection of the future forms in literature (e.g., Amaral 1920 and Said Ali 1969, sections 2.3.2. and 2.3.2.1.). The synthetic form is said to be used to express uncertainty, doubt, and

politeness in questions; hence, this hypothesis is examined in this study. Each token was coded for whether it occurred in statements, as in (37), or questions, as in (38):

(37) Daqui a dois dias não *sabe* (P) mais falar Português. (Pena 1842. D: Clemência: scene 11: 47)  
 ‘In two days he *won’t know* how to speak Portuguese anymore.’

(38) E eu como é que fico? *Vou virar* (GP) Teresinha Pinto? Deus me livre! (Buarque 1978. OM: Teresinha: scene 2: 58)  
 ‘How about myself? *Will* my name *be* Teresinha Pinto? God forbid!’

8 – *Presence and Type of adverbial specification*: Soares Barboza (1852, section 2.3.2.2.) comments that the synthetic future conveys an undetermined future. Thomas (1969), Cunha & Cintra (1984/ 1999) and Van Achter et al. (1996) claim that specific future time is indicated by the use of the present form with a temporal adverb (cf. section 2.3.3.3.). If this is the case, then one would expect specific time adverbs to favour present tense forms and absence of adverbial specification or non-specific adverbs to favour the synthetic form. In order to test this hypothesis, the data was coded for absence of adverbial specification, as in (39), for non-specific adverbials, as in (40), and for specific adverbs, as in (41):

(39) Não tou pagando nada, então não *vou estudar* (GP) não. Num precisa. (1992. Salvador: A: 24)  
 ‘I’m not paying anything, so I’m not *going to study*. It’s not necessary.’

(40) Periandro, *logo falaremos* (SF), não te ausentes. (Silva 1737. VE: Filena: first part, scene 2: 39)  
 ‘Periandro, we *will talk soon*. Don’t leave.’

- (41) Tá certo. Hoje eu faço vista grossa, mas **amanhã** eu te *caço* (P), viu? (Buarque 1978. OM: Max: scene 2: 57)  
 ‘Ok. Today I don’t care, but tomorrow I *will catch* you, see?’

9 – *Type of clause*: The type of clause in which the future forms appear has not been considered relevant by grammarians. However, one of the distinctions claimed as shaping the usage of *haver* + infinitive periphrasis is its related meaning to the Latin future participle (e.g., Barros 1540: 93-94; Leão 1608/ 1983: 298). As detailed in section 2.3.1.2., the active and passive Latin future participles correspond to relative clauses in Portuguese. In order to determine whether there was any special connection between relative clauses and the form *haver* + infinitive, tokens were coded to indicate whether they occurred in an independent clause (42), a subordinate clause (43) or a relative clause (44):

- (42) Eu *salvarei* (SF), mas debaixo de certas condições. (Pena 1853: N: Carlos: second act, scene 5: 96)  
 ‘I *will save* you, but with some conditions.’
- (43) Então diga ao teu marido que nós não *vamos precisar* (GP) do dinheiro dele, não. (Buarque 1978: OM: Terezinha: scene 3: 74)  
 ‘Tell your husband that we’re not *going to need* his money.’
- (44) Tolerância demais das autoridades que *vão permitir* (GP) essa imoralidade. (Gomes 1963. BH: Padre: first act: frame 8: 446)  
 ‘It’s too much tolerance by the authorities who *are going to allow* this immorality.’

10 – *Contingency*: It has often been noted that the synthetic future form expresses notions of doubt, probability and possibility (e.g. Amaral 1920; Bechara 1968; Cunha & Cintra 1984/ 1999; Kury 1989; Said Ali 1969; Rossi 1945; Thomas 1974; Van Achter et al. 1996; Arruda 2000, in section 2.3.3.1.). These notions have also been assigned to the *haver* + infinitive periphrasis (Barreto, 1944 and Thomas 1969, in section 2.3.3.2). In addition, Thomas (1969: 116) claims that the present form is used in future contexts for promises,

as in: *Se você me ajudar, eu lhe pago bem*. ‘If you help me, I’ll pay you well’. It seems then that a distinction is made between those events which can be assumed to have a high likelihood of occurring, and those which are less probable or contingent on other events. Following Fleischman (1982), Poplack and Turpin (1999), I coded contingent events as events whose realization is dependent upon the fulfillment of a condition, as in (45). Since the grammarians confer a hypothetical nuance to all forms, with the exception of *ir* + infinitive periphrasis, the working hypothesis is that this periphrasis will be associated to assumed events and actions in the language, as in (46):

(45) *Se eu fugir, eles me alcançam* (P) em dois tempos. E aí vão me manter dopado o resto da vida, (Labaki 1999: B: Ricardo: scene 4: 65)

‘If I run away, they *get* me very fast. Then they’re going to keep me on drugs for the rest of my life.’

(46) *Ninguém vai saber* (GP). Você sabe. A imprensa só publica o que deixam publicar. (Labaki 1999: B: Verônica: scene 4: 59)

‘Nobody *is going to know*. You know. The press only shows what they are allowed to show.’

11 - *Temporal distance*: Proximity has been frequently associated with the present in future contexts (e.g. Silva Junior 1887, in section 2.3.2.2.; Cunha & Cintra 1984/ 1999; Said Ali 1969; Silveira Bueno 1968, in section 2.3.3.3.) and with the *ir* + infinitive periphrasis (Cunha & Cintra 1984/ 1999; Kury 1986; Said Ali 1969; Thomas 1969; Arruda 2000, in section 2.3.3.4.). Furthermore, the association of *ir* + infinitive to proximate future is also made in French and Spanish (cf. Poplack and Turpin 1999; Bentivoglio and Sedano 1992, respectively). Taking these claims into consideration, the data are examined according to temporal distance between the speech time and the future event or action. The hypothesis is that proximal events and actions co-occur with the present and the *ir* + infinitive in Portuguese. Proximal event/ actions are defined in this

study as those that will occur in the same day of the utterance, as in (47); while distal event/ actions are those occurring in any other future time, as in (48):

(47) “Tá bom, num *vou ler* (GP) não.” Aí no meio do caminho, eu li, né. (1992. Érica: B: 13)  
 “‘All right. I’m not *going to read* it.’ But in the street, I read it.’

(48) Ano que vem eu *vou tirar* (GP) a Cleide dessa escola e colocar na escola pública, do governo. (1992. Vânia: B: 30)  
 ‘Next year, I’m *going to transfer* Cleide from this school to a public school, of the government.’

Three other hypotheses were postulated for testing against the data. The first one is related to the difference in the usage of the variants throughout the centuries, the second one is associated with the data source of the twentieth century, and the last is related to the speakers’ level of education:

12 – *Century of usage*: Not all four variants share equally in the partitioning of the future temporal reference sector in Portuguese, nor has the proportion of each remained consistent through out history. Coding individual tokens for the century of origin makes it possible to track changes in the use of variants throughout history and to compare the actual usage of the time to the description of that usage in contemporary grammars. So, for example, early grammar books did not acknowledge the usage of the present and the *ir* + infinitive periphrastic future forms. The present form was first mentioned in the nineteenth century, while the *ir* + infinitive periphrasis usage was first mentioned in 1945. The working hypothesis is that both variants were used before being legitimized by grammarians, and that only as their usage outstripped that of other variants did they attract the notice of prescriptive grammarians. Their usage in these centuries, in which they were

recognized in prescriptive literature, surpassed the usage of the other variants in the language.

13 – *Medium*: The majority of grammarians have claimed that synthetic variant is the future form *per se* (cf. Table 2.2). It is also the form mostly used in the formal written variety of Brazilian Portuguese (cf. Santos 1997, in section 4.3.). As stated in chapter 3, the plays selected for this study are expected to be representative of the spoken language, or at least, to be the closest written literary production. The working hypothesis of this study is that both *writing* and *speech* will be similar regarding the usage of the variants. This hypothesis can only be statistically tested against the twentieth century data because of the lack of spoken data from other centuries. In order to test it, each token was coded for whether it was collected from the written or the oral corpus in the twentieth century.

14 – *Level of Education*: This study's spoken data were gathered from Brazilian Speakers whose educational levels ranged from elementary to university (cf. section 3.3.2.3. for details). Recall that prescriptive literature established the synthetic form as being Portuguese's future form (cf. section 2.3.). Presumably, the longer a speaker is under the influence of formal education, the more he is likely to incorporate prescriptive language rules. If prescriptivism has echoed successfully in Brazilian schools, then the level of education should be a significant factor favouring the selection of the synthetic future. In that case, speakers with university level education will have a higher level of synthetic future form usage.

#### 4.5. Some Considerations on the Groups of Factors

All factor groups named in section 4.4, with the exception of medium and level of education<sup>26</sup>, were examined by means of the GoldVarb2 (Rand and Sankoff 1990). However, the following considerations regarding the factors should be made: some factors displayed in section 4.4 were scarcely represented in the data. Moreover, they frequently followed the same distributional pattern of another factor in the same factor group. As usual in such cases in variationist analyses, the factors were clustered or completely removed. Besides that, the factor groups of grammatical person and animacy showed a high degree of interaction. The cross-tabulation of these two factor groups reveal that animate and inanimate subjects are unevenly spread across the grammatical persons in the data. Table 4.3 depicts the distribution of the data:

*Table 4.3: Cross-Tabulation of the Grammatical Person and Animacy Group of Factors.*<sup>27</sup>

	<i>Human</i>	<i>Animate</i>	<i>Inanimate</i>	Total
1 <sup>st</sup> person	1024	5	—	1029
2 <sup>nd</sup> person	375	3	—	378
3 <sup>rd</sup> person	571	14	344	931
Total	1970	22	344	2338

Note that almost all 1<sup>st</sup> and 2<sup>nd</sup> grammatical persons are also human. Furthermore, about 85% of the data correspond to human subjects. Only in the 3<sup>rd</sup> grammatical person is the data balanced between human and inanimate factors; [+ animate, - human] subjects are scarcely represented in the corpus (only 22 tokens). This pattern is to be expected in data

<sup>26</sup> These factor groups were only applied to twentieth century data. They are examined in chapter 6.

<sup>27</sup> A total of 27 tokens occurred with clauses without subjects. They were not included in this Table.

representing spoken style. People tend to talk about themselves, their relatives, friends and acquaintances. It is due to chance that [+ animate, - human] elements are introduced in dialogues. Due to this distribution, I collapsed these two groups of factors. The factors are the followings: 1<sup>st</sup> animate person, 2<sup>nd</sup> animate person, 3<sup>rd</sup> animate person and 3<sup>rd</sup> inanimate person.

Finally, in a preliminary quantitative analysis, both *verbs of motion* and *monosyllabic verbs* were selected by the program as significant to the realization of the present variant. The possibility of interaction between these factor groups made it advisable to ensure that this was not merely derived from some lexical effect. Indeed, the distribution of the main verb *ir* 'to go' in future temporal reference contexts was revealing. The data is presented in Table 4.4. This Table shows the high correlation between the verb of motion 'to go' and the present variant occurrences in the Portuguese language. 73% of the total number of 'to go' usages as the main verb are expressed by the present in future contexts. Note in particular how the main verb 'to go' selects the present variant at the expense of the synthetic variant beginning in the eighteenth century. Indeed, in the twentieth century the present was selected 96% of the time the main verb 'to go' was used in future contexts. Synchronically, this choice can be considered as virtually categorical.

**Table 4.4: Distribution of the main verb *ir* ‘to go’ in future temporal reference contexts in the Portuguese corpus by variants.**

	<i>Synthetic future</i>	<i>Haver + infinitive</i>	<i>Present</i>	<i>Ir + infinitive</i>	Total N
	N	N	N	N	
16 <sup>th</sup> century	16	10	07	—	33
18 <sup>th</sup> century	01	03	09	—	13
19 <sup>th</sup> century	08	—	18	—	26
20 <sup>th</sup> century	01	—	78	02	81
Total N	26 17%	13 8%	112 73%	02 1%	153

From this it was clear that the lexical effect of the verb ‘to go’ was skewing the statistical results, except in the earliest periods where such an effect was not evident. Tokens of ‘to go’ were excluded from the analysis after the eighteenth century. As well, the factor group for number of syllables was excluded because of its dependence, and therefore interaction with, lexical identity.

The question that arose from this finding was whether the present variability could be attributed to other lexical effects of verbs of motion. Therefore, I decided to verify all other verbs that are used with this variant in the corpus. Thus, I listed all occurrences of the present variant with their corresponding verbs. The inventory is in Table 4.5.:

**Table 4.5: Occurrence and distribution of verbs with the present variant in the corpus. (N = number of occurrences)**

16 <sup>th</sup> century		18 <sup>th</sup> century		19 <sup>th</sup> century		20 <sup>th</sup> century			
						Written		Spoken	
	N		N		N		N		N
Ir go	7	Apanhar <i>get</i>	1	Acompanhar <i>accompany</i>	2	Acertar <i>get it right</i>	2	Acabar <i>finish</i>	1
Levar <i>take</i>	1	Dar <i>give</i>	1	Apanhar <i>get</i>	1	Alcançar <i>reach</i>	1	Acontecer <i>happen</i>	1
Voltar <i>come back</i>	2	Esperar <i>wait</i>	1	Border <i>knit</i>	1	Apelar <i>cheat</i>	1	Adiantar <i>ahead</i>	2
		Fazer <i>do</i>	1	Casar <i>marry</i>	1	Aproveitar <i>take advantage of</i>	1	Aposentar <i>retire</i>	1
		Ir go	9	Chamar <i>call</i>	1	Arranjar <i>arrange</i>	1	Aprender <i>learn</i>	2
		Levantar <i>get up</i>	1	Comprar <i>buy</i>	1	Assumir <i>take over</i>	1	Chegar <i>come out</i>	3
		Mandar <i>give order</i>	1	Correr <i>run</i>	1	Botar <i>put</i>	1	Colar <i>glue</i>	1
		Partir <i>leave</i>	1	Dar <i>give</i>	8	Caçar <i>hunt</i>	1	Começar <i>begin</i>	1
		Ter <i>have</i>	1	Deixar <i>leave</i>	1	Chegar <i>come out</i>	2	Comprar <i>buy</i>	1
		Tornar <i>turn</i>	1	Despachar <i>send</i>	1	Chorar <i>cry</i>	1	Conseguir <i>obtain</i>	1
		Vir <i>come</i>	1	Espalhar <i>spread</i>	1	Completar <i>accomplish</i>	1	Dar <i>give</i>	3
				Estar <i>be</i>	2	Continuar <i>continue</i>	1	Endoidar <i>get mad</i>	1
				Fazer <i>do</i>	4	Cuidar <i>take care</i>	1	Fazer <i>do</i>	1
				Ficar <i>stay</i>	1	Dar <i>give</i>	3	Ficar <i>stay</i>	2
				Ir go	18	Descontar <i>discount</i>	1	Formar <i>graduate</i>	1
				Largar <i>leave</i>	1	Embarcar <i>get in</i>	3	Ir go	57
				Levar <i>take</i>	2	Encaixar <i>fit</i>	1	Levar <i>take</i>	1
				Mandar <i>give orders</i>	1	Enterrar <i>bury</i>	1	Ligar <i>call</i>	1
				Matar <i>kill</i>	3	Escrever <i>write</i>	1	Morrer <i>die</i>	3
				Mostrar <i>show</i>	1	Estar <i>be</i>	3	Precisar <i>need</i>	1
				Mudar <i>move</i>	1	Explicar <i>explain</i>	1	Processar <i>process</i>	1
				Partir <i>leave</i>	1	Falar <i>speak</i>	3	Saber <i>know</i>	1
				Perder <i>lose</i>	1	Fazer <i>do</i>	2	Ser <i>be</i>	3
				Principiar <i>begin</i>	1	Ficar <i>stay</i>	2	Ter <i>have</i>	7
				Saber <i>know</i>	1	Formar <i>graduate</i>	1	Tirar <i>take off</i>	1
				Tardar <i>get late</i>	1	Ir go	21	Vir <i>come</i>	4
				Voltar <i>come back</i>	8	Morrer <i>die</i>	1	Voltar <i>come back</i>	2
						Oferecer <i>offer</i>	1		
						Pagar <i>pay</i>	1		
						Passar <i>pass/ become</i>	3		
						Perder <i>lose</i>	3		
						Precisar <i>need</i>	1		
						Receber <i>get</i>	1		
						Recrutar <i>assemble</i>	1		
						Reunir <i>meet</i>	1		
						Sair <i>leave</i>	3		
						Ser <i>be</i>	4		
						Ter <i>have</i>	4		
						Vender <i>sell</i>	1		
						Vir <i>come</i>	2		
						Voltar <i>come back</i>	8		

Table 4.5 clearly shows that the verb “to go” is unique in its strong lexical preference for the present variant. Nor does there appear to be any general effect of verbs of motion or stativity. Considering this distribution, the statistical analyses will retain all other verbs. From the total of 292 occurrences of the present variant in the future sector, 180 (62%) were realized by other verbs than ‘to go’.

In the remaining chapters, the results of the quantitative analyses regarding the selection of the variants in the future sector will be presented.

# Chapter 5 *Results*

---

In this chapter, I focus on the multivariate analyses of the future temporal reference variable in Portuguese. I first present some considerations on the distribution of the future variants. As a result, the analyses of the corpus by centuries are carried out separately. The results of the combined first two centuries of data, namely the sixteenth and the eighteenth centuries, as well as the results of the nineteenth century are provided here.

## 5.1. Overall Distribution of the Variants

Table 5.1 depicts the overall distribution of future temporal reference variants for all time periods.

---

*Table 5.1: Overall distribution of the variants of the future temporal reference in Portuguese.*

---

Variants	%	Number of tokens
<i>Ir</i> + infinitive Periphrasis	45	1077
Synthetic Future	30	693
<i>Haver</i> + infinitive Periphrasis	13	303
Present	12	292
<b>Total</b>		<b>2365</b>

---

Note that the *ir* + infinitive periphrasis is the dominant form, accounting for nearly half of all eligible contexts. The synthetic form is also well represented with an overall rate of 30%, followed by *haver* + infinitive and the present at rates of 13% and 12%, respectively. This distribution diverges sharply from the analyses of formal spoken and

written Portuguese presented by Baleeiro (1988) and Santos (1997), summarized in chapter 4 (cf. section 4.3). In order to form a clearer picture of how this distribution is patterned and to assess whether change has taken place, it was determined to examine the individual distribution of variants in each corpus by century.

## 5.2. Distribution of Variants by Century

An overall distribution of variants across centuries assumes that the linguistic and extra-linguistic factors conditioning the variation have remained stable and uniform over time. It is possible, however, that a number of distinct competitions could have emerged throughout the centuries with different variants participating to a greater or lesser degree. Indeed, the distribution of variants varies considerably from century to century, as shown in Table 5.2.:

**Table 5.2: Distribution of the Variants of the Future Temporal Reference in Portuguese by Century (all tokens)**

Century	<i>Synthetic future</i>		<i>Haver + infinitive</i>		<i>Present</i>		<i>Ir + infinitive</i>		Total	
	N	%	N	%	N	%	N	%		
16 <sup>th</sup>	198	66	91	30	10	3	3	1	302	
18 <sup>th</sup>	169	57	103	35	19	6	5	2	296	
19 <sup>th</sup>	276	53	104	20	66	13	72	14	518	
20 <sup>th</sup>	Written	46	9	5	1	93	18	384	73	528
	Spoken	4	1	--	--	104	14	613	85	721
Total	693	30	303	13	292	12	1077	45	2365	

Table 5.2 clearly shows a change from the earlier dominance of the synthetic and *haver + infinitive* variants to the increasing use of the present and *ir + infinitive* variants in the current language. The two earliest periods, the sixteenth and eighteenth centuries, show a very similar pattern of distribution: in both periods the synthetic and *haver + infinitive*

variants account for more than 90% of the future temporal sector. In the nineteenth century, however, the present and *ir* + infinitive variants begin to gain ground, spiking from 6% and 2% to 13% and 14%, respectively. By the twentieth century, the hierarchy has inverted itself and it is now the present and *ir* + infinitive variants that account for a startling 99% of the future sector. As dramatic as this change seems, it is important to rule out the possibility that the change in overall distribution does not represent actual language change, but a difference in distributions between spoken and written language, as the twentieth century is the only period to include spoken data. A comparison of the spoken and written data in the twentieth century corpus reveals the same tendencies for the use of each variant (Table 5.2), confirming that the plays selected do in fact represent speech, although the rates lag behind those of the spoken language. Even if the spoken data were excluded, the written data alone would suffice to demonstrate the reversal of rates in use of the variants.

Insofar as the differences in rate cannot be attributed to the differences between speech and writing, it seems that the distribution tendencies of future variants represent different stages in the development of the grammar. Individual multivariate analyses of each period represented should not only uncover the individual constraint hierarchies operational at any given time, but make it possible to trace how the ‘new’ variants, i.e. the present and the periphrastic form *ir* + infinitive, came to rival and eventually dominate the traditional, established forms.

### 5.3. The Sixteenth and Eighteenth Centuries

In both the sixteenth and eighteenth centuries, expression of future temporal reference in Portuguese was dominated by two variants, namely the synthetic future and the *haver* + infinitive. The distributional pattern appears to have remained quite stable

throughout this period, therefore a multivariate analysis was run on the combined results, the distribution of which is depicted in Table 5.3:

**Table 5.3: Distribution of the Variants of the Future Temporal Reference in Portuguese after Combining the Sixteenth and the Eighteenth Century Data.**

	<i>Synthetic future</i>		<i>Haver + infinitive</i>		<i>Present</i>		<i>Ir + infinitive</i>		Total
	N	%	N	%	N	%	N	%	
<b>16<sup>th</sup> and 18<sup>th</sup> centuries</b>	367	61	194	33	29	5	8	1	598

Note that *ir* + infinitive periphrasis, as exemplified in (1), accounts for only 1% of the data.

(1) Deixe-o enforçar, que eu também *vou fazer* (GP) o mesmo. Arre com a cancaburrada da noitezinha! Olhem, não há coisa mais fiel, que o nariz; por isso lhe fedia o bafo a cevada; mas como tinha o nariz cego de amor, cuidei, que me cheirava a benjoim. (Silva 1734. VE: Esopo: second part: scene 6: 132)

‘Let him hang himself, for I *will do* the same. Damn the night! Look, there isn’t anything more trusting than the nose; that is why the breath smelled of barley; but since my nose was blind because of love, I convinced myself that it smelled of benzoin.’ [Situation: he kissed a man by mistake]

The extreme rarity of this variant is noteworthy. In the sixteenth century, there are only three occurrences of this periphrasis out of 302 tokens of future temporal reference. Two hundred years later, it still represents only five out of 296 tokens. Based on such scarcity, one must assume that the variant was either not used at that time or was underrepresented in the written language in spite of its informal nature (cf. chapter 3). Although no quantitative analysis exists, the equivalent form is found to have occurred in

both French and English in the seventeenth century (cf. Poplack and Turpin 1999: 134; Poplack and Tagliamonte 2001: 211). The very low rates discovered for the Portuguese use of this variant during this period may suggest that this periphrasis also had a low frequency of usage in these languages. Clearly, *ir* + infinitive is only incipient at this time.

With so few data it was not possible to include the *ir* + infinitive variant in a multivariate analysis. A closer examination of all eight tokens of this variant however, revealed that they all co-occurred with the following factors: affirmation, statements as opposed to questions and assumed, as opposed to contingent events. Seven of the eight tokens also occurred with stative verbs and human subjects. Thus, the linguistic structure that allowed the usage of the periphrastic variant *ir* + infinitive was basically the same: affirmative statements, with stative verbs, human subjects and assumed events. Moreover, the features of affirmative, statement, assumed events, and human subjects represent independently more than 80% of the data of the sixteenth and the eighteenth centuries. Thus, while the variant itself was extremely rare, the context that permitted this variant was very robust in the data, suggesting that this variant was first licensed in the presence of features that show a very high frequency in discourse. From this point of view, these features possibly configured the least salient linguistic contexts.

After the eight tokens of *ir* + infinitive were considered independently, a multivariate analysis was performed on the remaining 590 tokens. The results are presented in the following section.

### 5.3.1. *The Sixteenth and Eighteenth Century Variable Rule Analyses*

#### 5.3.1.1. *Factors*

Three independent variable rule analyses were carried out in order to determine which factors favoured the choice of the synthetic future, the periphrastic *haver* + infinitive variant, or the present variant in the future temporal sector during the first two centuries of the corpus.

Some considerations, however, need to be taken into account regarding the factor groups that were tested in these analyses. In section 4.5, it was mentioned that the main verb “to go” showed a strong lexical tendency to favour the present variant (cf. Table 4.4). However, this lexical effect had no evident effect on the distribution of variants in the earlier periods, so these tokens were included in the multivariate analysis of the sixteenth and eighteenth centuries only. Furthermore, the prescriptive notion that the *haver* + infinitive construction is restricted to relative clauses with future-oriented verbs, is not supported by the distribution of that variant, which makes up roughly a third of all contexts, regardless of clause type (32% of the main clause contexts, 32% of subordinate contexts, and 38% of relative clause contexts.) Since the relative/non-relative distinction did not appear to be relevant in this context and the number of tokens in subordinate and relative clauses was not robust (55 and 19, respectively), it was decided to collapse the relative and subordinate clause groups together to determine whether the effect of main versus subordinate clause was operational on these data. Finally, some factor groups were altered or excluded to avoid interactions. As displayed in Table 4.3, there was a strong tendency for certain subject types and animacy to co-occur, thus these groups were recombined to create a new factor group (cf. section 4.5). The factor group for number of syllables had to be excluded from the multivariate analyses due to a number of interactions with other

linguistic factors which made it impossible to distinguish any independent effect. All other linguistic factors were considered in the analysis of the future temporal sector in the sixteenth and eighteenth centuries (cf. section 4.4.).

#### *5.3.1.2. Results*

Table 5.4 shows the results of three independent variable rule analyses of the contribution of linguistic factors selected as significant to the choice of variants in the future temporal reference contexts in the sixteenth and eighteenth centuries. The magnitude of the effect of each factor in the choice of the variant is given by the range after each factor group in the Table. The factors that were not selected as significant by the stepwise multiple regression procedure are displayed at the bottom of the Table. The results are presented by variant.

**Table 5.4: Three variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic future, the Haver + infinitive periphrasis and the Present will appear in future temporal reference contexts in the sixteenth and eighteenth centuries in Portuguese.**

	<i>Synthetic future</i>	<i>Haver + infinitive</i>	<i>Present</i>
Total N (590)	367	194	29
Corrected Mean	.622	.329	.049
	% N	% N	% N
<b>Contingency</b>			
Contingent	.67 78 35	.31 18 08	
Assumed	.48 61 332	.52 34 186	
Range	19	21	
<b>Sentence Type</b>			
Statement	.53 65 322	.47 30 148	
Question	.36 47 45	.66 48 46	
Range	17	19	
<b>Stativity</b>			
Motion	.36 48 59		.87 17 21
Stative	.54 66 308		.38 02 08
Range	18		49
<b>Polarity</b>			
Affirmative	.51 64 340		
Negative	.37 49 27		
Range	14		
<b>Temporal Distance</b>			
Proximal			.72 10 23
Distal			.35 02 06
Range			37
<b>Grammatical Person/Animacy</b>			
1 <sup>st</sup> Animate			.68 07 18
2 <sup>nd</sup> Animate			.20 01 02
3 <sup>rd</sup> Animate			.41 03 04
3 <sup>rd</sup> Inanimate			.62 06 05
Range			48
Factor groups not selected as significant:			
Temporal Distance	X	X	
Grammatical Person/ Animacy	X	X	
Stativity		X	
Polarity		X	
Contingency			X
Sentence Type			X
Adverbial Specification	X	X	X
Presence of Clitics	X	X	X
Type of Clause	X	X	X

### 5.3.1.2.1. The Synthetic Variant

The results for the synthetic variant reveal that its occurrence in the future temporal sector was constrained by four groups of factors in the sixteenth and eighteenth centuries: *contingency*, *stativity*, *sentence type* and *polarity*<sup>28</sup>(cf. Table 5.4).

The factor group that contributes the most to the selection of this variant is *contingency*, as shown by its range (range = 19). Recall that contingent events/actions, as in (2), were those dependent upon circumstances, in other words, those more hypothetical in their future occurrences, while *assumed*, as in (3), were not indicated as dependent upon the fulfillment of a condition (cf. section 4.4.):

- (2) Não senhora; eu *virei* (SF) logo nessora, se m'eu lá não detiver. (Vicente 1509. AI: Moça: 382)  
'No mother; if I don't have to stay there, I *will come* soon, right at this time.'
- (3) Aqui *ficarei* (SF) esperando por esse Sol, que me anima. Ai amor...(Silva 1734. VE: Periandro: first part: scene 2: 39)  
'I *will stay* here waiting for the sun that liven me up. Oh love...'

Contingent contexts as in (2) favoured the occurrence of the synthetic future with the probability of .67 in the sixteenth and eighteenth centuries. This suggests that the synthetic future may be favoured in hypothetical contexts or to convey notions of doubt or possibility.

The next factor group selected as significant to the occurrence of the synthetic future is *stativity*. Note in Table 5.4 that this variant is used with both *stative verbs* and

---

<sup>28</sup> The cross-tabulation of both sentence type and polarity factor groups does not exhibit empty cells. However, the number of tokens in affirmative and statement contexts is greater than in the negative and questions. That is due to the configuration of the corpus. It not only happened in those two centuries, but also in the others.

*verbs of motion* in the language; however, *stative verbs*, as in (4), favour its selection with the probability of .54.

- (4) Senhor, eu o *servirei* (SF) e não quero outro senhor, (Vicente 1526. FA: Ourives: 183)  
 ‘Sir, I *will serve* you and I don’t want other master.’

Sentence type also has an influence on the selection of the synthetic variant, with statements favouring the synthetic future (as in 5) and questions disfavouring it.

- (5) Ela lhe *dará* (SF) o recado. (Silva 1734. VE: Periandro: first part: scene 2: 34)  
 ‘She *will give* you the message.’

The last factor to contribute to the selection of the synthetic variant in the future temporal context is *polarity*. Table 5.4 reveals that, although the synthetic future may also occur with negation, it is preferred in affirmative contexts in the sixteenth and eighteenth centuries, as in (6):

- (6) Periandro, logo *falaremos* (SF), não te ausentes. (Silva 1737. VE: Filena: first part, scene 2: 39)  
 ‘Periandro, we *will talk* soon. Don’t leave.’

Overall, the probability of having the synthetic future in future temporal contexts was .622, statistically demonstrating that the synthetic variant was the overall preferred choice for expressing the future temporal reference during the sixteenth and eighteenth centuries in Portuguese. This was not a random choice, however, but was conditioned by the presence of specific favouring linguistic factors: *contingent events/actions*, *stative verbs* and *affirmative statements*. Of these effects, contingency was the most significant.

If we compare these results with grammarians' prescriptions (cf. Table 2.2), it is clear that none of these conditioning factors were identified by grammarians of the time. The only meaning attributed to the synthetic future was that of strict future reference, without additional nuance. During the sixteenth and eighteenth centuries the prescriptive literature did not associate any specific meaning or linguistic context to the usage of this variant. Thus, despite the grammarians' prescriptions, Table 5.4 clearly shows that the synthetic future did share the temporal reference sector with other variants and that the selection was linguistically conditioned as far back as the sixteenth century.

#### 5.3.1.2.2. *The Periphrastic Haver + infinitive Variant*

The analysis selected two factors as significant in the selection of the periphrastic *haver* + infinitive variant in the sixteenth and eighteenth centuries: *contingency* and *sentence type* (cf. Table 5.4).

In the *contingency* factor group, *assumed events/actions* showed the greatest preference for *haver* + infinitive in future context, as in the example in (7). As Table 5.4 shows, this variant is preferred where the validity of the future eventuality is assumed, i.e. when the future context is characterized as being less hypothetical. The results of the contingency factor group for both the synthetic future and *haver* + infinitive variants show that the choice between these two variants was highly influenced by this factor (cf. Table 5.4) with contingent events favouring the synthetic future (as noted above) and assumed ones favouring *haver* periphrasis.

- (7) Não vo-lo *hei de negar* (HP): *fazei-me uma petição*. (Vicente 1526. FA: Fidalgo: 46: 176)  
'I *won't deny* it to you: make your request.'

The next factor group contributing to the selection of *haver* + infinitive is *sentence type*. Table 5.4 shows that questions favour *haver* + infinitive with a probability of .66, as in (8).

- (8) Pois aonde o *hei de pôr* (HP)? (Silva 1737. VE: Semicúpio: scene 2: 54)  
 'Where *will* I *put* it?

Thus the data indicate clear criteria selecting between variants with contingent events and statements favouring the synthetic future and assumed events and questions favouring *haver* + infinitive. How do these results correspond to the picture presented by prescriptive grammarians in these centuries? Recall that grammarians did not associate the use of the synthetic future with any specific meaning or linguistic context in these centuries. In chapter 2, this fact was explained in detail. Since the synthetic future was viewed as the default future form by grammarians, the periphrastic future *haver* + infinitive was considered to express meanings other than futurity (cf. Table 2.2).

The following meanings and linguistic context were associated to this periphrasis: obligation, speaker's intention/determination, proximity of the future eventuality and its usage in relative clause. As explained in chapter 2, some meanings attached to the future forms could not be operationalized; thus, they cannot be empirically tested. At the bottom part of Table 5.4, we see that the factor groups of (i) temporal distance, which tested the effect of proximity of the future eventuality/action, and (ii) type of clause, which tested the effect of relative/subordinated clauses, were not selected as statistically significant to the selection of the periphrastic *haver* + infinitive variant in the future temporal context. Thus, these prescriptive associations are not supported in this study.

At the same time, grammarians appear to have failed to recognize the preference for *haver* + infinitive in assumed contexts and in questions. These being the two linguistic factors that in fact contribute to the choice of this periphrasis in the sixteenth and eighteenth centuries (cf. Table 2.2).

Summing up, Table 5.4 shows that at this time, 95 %, i.e. almost all reference to future is made up of the synthetic future and the periphrastic *haver* + infinitive variant. Although the synthetic future is the more frequent variant (62%), the periphrastic is also represented, constrained by the semantic factor of contingency of the event, and the syntactic factor of sentence type in the following ways; (i) events viewed as contingent upon circumstances favour the synthetic future, while the periphrastic form *haver* + infinitive is preferred for assumed future contexts; (ii) statements favour the occurrence of the synthetic future, while questions favour the selection of *haver* + infinitive variant.

#### 5.3.1.2.3. *The Present Variant*

The present variant is exceedingly rare at this time, accounting for less than 5% of all future temporal reference. Nonetheless, its occurrence is still regularly conditioned. The present is favoured by three factors, neither of which is relevant to the choice of either of the other variants. These factors are: *verbs of motion*, *1<sup>st</sup> and 3<sup>rd</sup> inanimate grammatical person* and *proximal events and actions*.

The first factor in order of significance is *stativity* (range of 49). In Table 5.4, we see that this group is divided into two factors: stative verbs and verbs of motion. The present variant is favoured by verbs of motion with the probability of .87. As we saw in chapter 4 (Table 4.5), out of the 21 occurrences of the present variant with verbs of motion,

16 (76%) occurred with the verb 'to go'. Thus, the effect of verbs of motion may be better interpreted as a lexical effect of single verb on the data of the sixteenth and eighteenth centuries.

The next group of factor selected as statistically significant to the choice of the present variant is *grammatical person/animacy* (range of 48). Table 5.4 shows that the 1<sup>st</sup> animate and the 3<sup>rd</sup> inanimate persons favour the choice of the present variant in future contexts with the probability of .68 and .62, respectively. The effect of 1<sup>st</sup> animate subjects could be explained by the strong correlation between such subject types and the kinds of meanings that are commonly associated with the present variant: intention and certainty. Both meanings are related to the agentivity of the subject. Thus, intention and certainty of the future event occurrence is higher if the person who utters the sentence will perform it (cf. section 4.4.). However, this does not explain the favouring effect of 3<sup>rd</sup> person inanimate subjects.

The last factor group that exhibits a significant effect on the present variant is *temporal distance* between speech time and future eventuality. Recall that this factor group was divided into two factors: *proximal* and *distal*. The data was originally divided into four temporal distances: today, this week, this year, all others. As the results did not show a systematic association of forms with these four categories, it was decided to collapse the data into two categories, proximal, i.e. events/actions occurring in the same day of the utterance, and distal, events/actions that will occur in any other future time (cf. section 4.4.). As shown in Table 5.4, in the sixteenth and eighteenth centuries, *proximal* future eventuality favours the selection of the present with the probability of .72.

Given that there were only 29 tokens with the present variant, there was some concern over possible interaction among the selected factors or an excess of empty cells. Cross-tabulations showed no empty cells; however, 19 out of the 23 tokens with proximal reference, and 12 out of the 18 tokens with 1<sup>st</sup> person were also verbs of motion (76% = ‘to go’), as exemplified in (9).

- (9) Vai tu, filho Joane,/ e dize que logo vou (P),/ que não faz tempo que cá estou. (Vicente 1512. VH: Velho: 83)  
 ‘Go, my son Joane,/ and tell (them) that I *will go* soon, / because I’ve been here for not so long.’

This association is likely the cause of the selection of both proximal and 1<sup>st</sup> person factors as having favoured the present. Thus, the apparent effects of 1<sup>st</sup> person and proximity appear to be epiphenomena of the association between the verb of motion ‘to go’ and the present tense. The results show that this association is as old as the first attestation of the present in future temporal reference contexts.

When comparing these results with grammarians’ remarks (cf. Table 2.2), we notice that from the sixteenth to the eighteenth centuries grammarians did not acknowledge the use of the present variant in the future temporal context, and in fact, it is very rare in the plays studied (N = 29). Nevertheless, already at that time the present variant showed a high propensity of being associated with verbs of motion, particularly to the verb ‘to go’, in Portuguese.

The next section presents the results of the nineteenth century. This will demonstrate how the present variant and the introduction of the ‘new’ periphrastic form *ir* + infinitive to the future sector contribute to the eventual demise of the synthetic future and *haver* + infinitive, respectively, in Brazilian Portuguese.

#### 5.4. The Nineteenth Century

This section examines the usage of the variant forms of future temporal reference in nineteenth-century Portuguese. As detailed in chapter 3, in contrast to preceding centuries, the data used here could be drawn from Brazilian popular plays. The Future Temporal Reference is expressed by four variants in this century: the *synthetic future*, *the present*, *haver* + infinitive and *ir* + infinitive periphrases.

As shown in Table 5.2, the nineteenth century data set is distinguished from the previous ones by an increase in the usage of the ‘new’ variants, namely the present and *ir* + infinitive, and a decrease in the occurrences of *haver* + infinitive periphrasis in the variable context. The ‘new’ variants account together for 27% of the expressions of future temporal reference in this century, in contrast to 6% in the preceding centuries (cf. Table 5.3). The synthetic variant is still the predominant form for the expression of the future (53%); but the data do not support the weight given to this variant as the default form by prescriptive grammars in this century (cf. chapter 2).

As for previous centuries, hypotheses were operationalized as factors on the data. All factor groups considered in the analyses of the previous centuries were duplicated in the analyses of the nineteenth century data. However, remember that the main verb *ir* ‘to

go' did not co-occur with the periphrases in the nineteenth century (cf. Table 4.4); thus the tokens with the main verb 'to go' were not included in the quantitative analysis.

Table 5.5 displays the results of the four variable rule analyses of the contribution of linguistic factors to the choice of each variant in future temporal reference contexts in the nineteenth century Portuguese.

**Table 5.5: Four variable rule analyses of the contribution of factors selected as significant to the probability that Synthetic future, Haver + infinitive, Present and Ir + infinitive will appear in future temporal reference contexts in the nineteenth century in Brazilian Portuguese. Without the main verb ir 'to go'.**

	<i>Synthetic future</i>	<i>Haver + infinitive</i>	<i>Present</i>	<i>Ir + infinitive</i>
Number (492)	268	104	48	72
Corrected Mean	.545	.211	.098	.146
	% N	% N	% N	% N
<b>Adverbial Specification</b>				
Non-specific	.69 72 58		.45 10 08	.15 02 02
Specific	.41 42 18		.88 33 14	.28 09 04
No adverbial	.47 52 192		.45 07 26	.62 18 66
Range	28		43	47
<b>Stativity</b>				
Stative	.52 56 249		.44 07 29	
Motion	.31 37 19		.89 37 19	
Range	21		45	
<b>Temporal Distance</b>				
Distal	.56 61 206	.55 24 81		.35 07 24
Proximal	.37 41 62	.40 15 23		.79 31 48
Range	19	15		44
<b>Contingency</b>				
Assumed		.52 22 101	.47 09 40	
Contingent		.27 09 03	.83 24 08	
Range		25	36	
<b>Polarity</b>				
Affirmative				.57 16 71
Negative				.09 02 01
Range				48
<b>Factor groups not selected as significant:</b>				
Adverbial Specification		X		
Stativity		X		X
Temporal Distance			X	
Contingency		X		X
Polarity	X	X	X	
Grammatical Person/ Animacy	X	X	X	X
Sentence Type	X	X	X	X
Presence of Clitics	X	X	X	X
Type of Clause	X	X	X	X

### 5.4.1. The Synthetic Variant

In the nineteenth century, the synthetic future was constrained by three groups of factors: *adverbial specification*, *stativity* and *temporal distance*.

As Table 5.5 indicates, the major constraint influencing the selection of the synthetic future in future contexts is *adverbial specification*. This factor group is subdivided into three factors: contexts of non-specific adverbial, as in (10); specific adverbials, as in (11); and no adverbial complement, as in (12)<sup>29</sup> (cf. section 4.4).

(10) Que impertinência! **Logo** conversaremos. (Pena 1842. D: Mariquinha: scene 2: 25)  
'Impertinent! [We] *will talk soon*.'

(11) O livro que te prometi, *mando* (P) **amanhã**. (Pena 1842. D: Cecília: scene 11: 50)  
'The book that I promised to you, I *will send* [it] **tomorrow**.'

(12) *Vou transmitir* (GP) essa ordem ao porteiro, porque eu posso não estar na ocasião.  
(Azevedo 1897. CF: Gerente: scene 6: 23)  
'[I] *'m going to give* this message to the doorman because I may not be here in the occasion.'

First, note that the synthetic future co-occurs with all three types of adverbial specification (cf. Table 5.5). However, non-specific adverbs, as in (10), favour its selection in future contexts with a probability of .69, whereas a specific adverbial, or no adverbial modification slightly disfavour the synthetic variant.

The next group of factor selected as having a highly significant effect on the occurrence of the synthetic future variant is *stativity*. As presented in the table, in the nineteenth century stative verbs favoured the selection of the synthetic future. This is the same direction of effect witnessed in the sixteenth and eighteenth centuries (cf. Table 5.4),

---

<sup>29</sup> I replicate here the examples presented in section 4.4., in which this factor group is explained in detail.

with the synthetic future being favoured with stative verbs, and disfavoured with verbs of motion. The range, or magnitude of this effect, however, has increased to 21 in that century.

The last factor group constraining the selection of the synthetic variant is *temporal distance* of the future eventuality. This factor group was subdivided into proximal and distal temporal distances. Tokens were coded for proximal distance when the future eventuality would occur in the same day, any other distance in the future time was coded as distal (cf. section 4.4). Table 5.5 shows that the *distal* factor significantly favours the choice of the synthetic variant. During the nineteenth century, the synthetic future was preferred for future events/actions that would not occur in the same day of the utterance.

How are these results related to the usage of the synthetic future in previous centuries? As already mentioned, the influence of *stativity* on the selection of the synthetic variant shows a similar pattern across the centuries. In the analyses of the sixteenth and eighteenth centuries as well as in the analyses of the nineteenth century, the synthetic variant is favoured by stative verbs. The other favouring factors in the sixteenth and eighteenth centuries, namely affirmative statements and contingent eventuality, no longer appear to be operational in nineteenth century Portuguese. On the other hand, now the synthetic future is highly favoured by adverbial specification and in a lesser extent by temporal distance (cf. Table 5.5.)

The findings of the sixteenth and eighteenth century and the nineteenth century quantitative analyses with regard to the synthetic future may, at first, seem to be inconsistent. However, recall that there were a greater number of variants expressing the future temporal reference in the nineteenth century. Thus, the synthetic future lost ground

in the language (54%) at the expense of the emergent forms. So, for example, the usage of the present variant was expanded to contingent context, which previously favoured the synthetic future occurrence. As the usage of a variant recedes in the future sector, some favouring contexts are taken by other variants in the following centuries. In section 5.5, an overall comparison of contexts by century will be presented.

The final consideration is whether the results in Table 5.5 reflect the prescriptive remarks for the usage of the synthetic future in the nineteenth century. During that century, the grammarians viewed the synthetic future as expressing only future time, with the exception of Soares Barboza (1852). According to this author, the synthetic future expressed positive statements and undetermined future time (cf. Table 2.2). As presented in section 4.4., it was inferred that undetermined future time meant contexts of a non-specific adverbial. Therefore, the empirical results in Table 5.5 corroborate Soares Barboza's claim, and, consequently, confirm my working hypothesis. This is the only possible correspondence between prescription and usage with regard to the synthetic future in this century.

#### 5.4.2. *The Periphrastic Haver + infinitive Variant*

In this section, I turn my attention to the results for the periphrastic *haver* + infinitive variant in the nineteenth-century plays. Table 5.5 shows that this periphrasis is constrained by two factor groups: *contingency* and *temporal distance*.

The factor group of *contingency* is the one which contributed the most to the occurrence of this variant in the future sector in the nineteenth century. *Haver* + infinitive exhibits an association with *assumed future events or actions* (probability of .52). Since the

realization of assumed events/actions is not dependent of any condition, this result confers to this periphrasis a weak hypothetical value.

The second factor group constraining the selection of the periphrastic *haver* + infinitive variant is *temporal distance*. Events coded as referring to *distal* future promoted the selection of this periphrasis in future contexts. Table 5.5 shows that distal future favours this periphrasis with the probability of .55. However, this factor group exerts a weaker effect on this variant choice than contingency, as expressed by their ranges, 15 and 25, respectively.

When we compare these results with the results of the previous centuries, we note that the temporal distance effect becomes statistically significant on the choice of *haver* + infinitive only in the nineteenth century.

On the other hand, recall that in the sixteenth and eighteenth centuries, the factor group of *contingency* showed the most robust effect on the realization of *haver* + infinitive, and, as Table 5.5 reveals, it continues to be the most important in the nineteenth century. In the analysis of the previous centuries, this factor distinguished the usage of *haver* + infinitive from the usage of the synthetic future. In the nineteenth century, however, it competes with the present variant, which is strongly favoured in contingent contexts (with a probability of .83). The results show that *haver* + infinitive has been consistently preferred for assumed events or states since the sixteenth century in Portuguese.

With regard to prescriptive remarks, in the nineteenth century *haver* + infinitive was associated to the meanings of possibility, firm resolution, necessity and intention, besides futurity (cf. Table 2.2). Firm resolution and intention could be related to the selection of *haver* + infinitive with assumed events; however, possibility and necessity are unlikely

related to this factor. Furthermore, none of these meanings can be correlated to the usage of this periphrasis with distal states and events. Thus, the results in Table 5.5 confirms that *haver* + infinitive was variably used in the future sector, and it was constrained by linguistic factors that were not acknowledged in nineteenth-century Portuguese grammars.

#### 5.4.3. *The Present Variant*

Three factor groups played a role in the selection of the present variant in the nineteenth century: *stativity*, *adverbial specification* and *contingency*.

*Stativity* exerts the highest effect on the selection of the present variant in future contexts. *Verbs of motion* favour the choice of the present with the probability of .89. Recall that in the nineteenth-century, the main verb ‘to go’ is not included in the analyses. Thus at this time the effect of verbs of motion on the selection of this variant is not interpreted as a lexical one.

An effect of almost equal magnitude is contributed by *adverbial specification*. In Table 5.5, the result obtained for this group of factors clearly demonstrates that the present variant is favoured with adverbs referring to a *specific time* in future contexts during the nineteenth century (probability of .88). When the present variant is used an overt time reference in the context does not make the tense marking redundant. On the contrary, the preference for the present variant in specific adverbial contexts seems to correspond to a disambiguating effect in the language, since of the four future variants the present is the only one whose usage could be understood as referring to another tense, i.e. present temporal reference instead of the future one.

The last selected factor group measured the probability of this variant being chosen in the future context when the eventuality is distinguished between contingent and assumed events/actions. Table 5.5 shows that *contingent* events favoured the present variant in the nineteenth century, which attributes to this variant a more hypothetical value.

If we examine the trajectory of the present variant from the sixteenth to the nineteenth century, we notice that the favouring effect of verbs of motion has been constant. Throughout the period studied, these verbs have favoured the present variant. Two additional factors favouring the present form in the nineteenth century are reported for the first time. The selection of contingency was particularly interesting. As observed in the previous section, assumed events have consistently favoured the periphrastic *haver* + infinitive through the centuries. The effect of contingent events, however, varied over time in the sense that they favoured the selection of the synthetic future in the sixteenth and eighteenth centuries and then favoured the selection of the present future in the nineteenth century. From the nineteenth century on, the synthetic future is no longer selected as significant in the expression of contingency. The present variant continues to be preferred for contingent events, while the synthetic future usage continues to recede in the future sector in Portuguese.

The findings of this analysis with respect to prescriptive comments indicate that the grammar books investigated did not represent adequately the usage of the present variant in the nineteenth century (cf. Table 2.2). First, recall that only in this century, was this variant even mentioned. In 1887, it was first cited as being used for proximal events or actions in the future. As we can see in Table 5.5, this remark does not correspond to any of the factors

favouring the choice of the present, namely, *verbs of motion, specific adverbs* and *contingent events*.

As most of the hypotheses pertaining to the rise of the present variant were based on twentieth century grammars, they will be presented after the discussion of the twentieth century results. In summary, the quantitative analysis confirms that the present variant is variably used in future contexts. Moreover, it shows that its usage was regularly constrained by linguistic factors. Verbs of motion, specific adverbial references and contingent events/actions comprised the preferred locus for the selection of the present in the Brazilian Portuguese language in the nineteenth century.

#### 5.4.4. *The Periphrastic Ir + infinitive Variant*

So far I have provided results for variants that occurred in the previous centuries. The focus of this section is on the periphrastic construction with *ir* + infinitive, that is, the form that currently dominates the future temporal reference contexts in Portuguese, as well as in French (Poplack and Turpin 1999) and to a lesser extent, English (Poplack and Tagliamonte 2000, 2001). In Portuguese, this form was introduced in the nineteenth century. Our analysis enables us to examine the constraints on its occurrence at this early time.

As mentioned in section 5.3, a prominent presence of this variant in the variable context was expected earlier in the Portuguese data. This expectation was based on two facts. First of all, the confirmation by prescriptive literature that *ir* + infinitive was used to express futurity in the middle of the twentieth century in Portuguese (cf. chapter 2). Based on the assumption that prescriptive comments regarding non-standard forms commonly

appeared when these forms were already well established, it was expected that the earliest usage of this periphrasis would probably be accessible in informal written Portuguese of the sixteenth or eighteenth centuries. Second, the occurrence of this form in seventeenth century French and English (cf. Poplack and Turpin 1999: 134; Poplack and Tagliamonte 2001: 211), made it reasonable to suppose that the form should have emerged in Portuguese around that time. However, the data revealed that *ir* + infinitive did not actually compete for the variable context until the nineteenth century, at least in the plays I have studied.

Despite this late appearance in the data, *ir* + infinitive had already surpassed the usage of the present in the variable context by this century. Table 5.5 shows that its occurrence was affected by three factors: *polarity*, *adverbial specification* and *temporal distance*.

The factor group of *polarity* shows the highest effect on the choice of the periphrastic *ir* + infinitive variant, as expressed by its range (48). *Affirmative* utterances favour the occurrence of this periphrasis with the probability of .57. Negative utterances, on the other hand, almost categorically reject *ir* + infinitive.

*Adverbial specification* also contributed a strong statistically significant effect to the occurrence (range = 47) of this variant, with *ir* + infinitive clearly being favoured in utterances with *no adverbial specification* (cf. Table 5.5). Note that this periphrasis entered the future variable context after it lost its original meaning, namely that of moving towards a goal/place in order to perform an action expressed by a verb in its infinitival form. Thus, in Portuguese the configuration of these utterances was typically formed by the verb *ir* 'to go' in the present tense combined with an infinitive, which represented an action to be performed, and adverbs of place. Furthermore, these adverbs could be placed before or

after the infinitives. Table 5.5 shows that temporal adverbials were not regularly added to this configuration when the future meaning started to be attached to this periphrasis, at least not at first.

An interesting finding of the quantitative analyses with regard to the adverbial specification is the pattern that emerged in the future sector in the nineteenth century. Table 5.5 reveals that the synthetic future, the present and *ir* + infinitive variants divide up the adverbial spectrum. Each variant is favoured by contexts with different adverbial specification: the synthetic future is favoured in contexts with non-specific adverbials, the present variant is highly favoured in contexts with specific adverbials, while *ir* + infinitive is highly favoured in contexts with no temporal adverbial complements. Thus, in the nineteenth century adverbial specification was not selected as statistically significant only to the choice of *haver* + infinitive.

The last selected factor was *temporal distance*. In Table 5.5, we can see that in the nineteenth century data, *ir* + infinitive was clearly associated with *proximity* in the future. If we examine the results displayed in Table 5.4 for the sixteenth and the eighteenth century data, we will notice that the established variants, the synthetic future and *haver* + infinitive, were not distinguished from each other in terms of temporal distance. Only the present variant was favoured in the proximal future context in the sixteenth and the eighteenth centuries. In the nineteenth century, this same context favoured *ir* + infinitive, i.e. the “new” variant which entered the future sector at that time (cf. Table 5.5). According to the quantitative results displayed in both tables, the present and *ir* + infinitive variants extended their dominion into the future sector from proximal future, i.e. this factor seems to

be one of the “open doors” for the admission of new variants in the future sector in Portuguese.

In summary, in the nineteenth century the periphrastic *ir* + infinitive variant is introduced as a variant of the future temporal reference in the plays under analysis. The findings of the quantitative analysis indicate that its occurrence is not random. This periphrasis is favoured in very specific linguistic contexts: (i) in affirmative utterances, (ii) with no adverbial specification, and (iii) in proximal future events/actions.

The next section presents a comparison of the results of the variable rule analyses. By means of comparison, the trajectory of the variants will be reviewed up to the nineteenth century. Hence, the direction of change will be pointed out in order to enable us to follow this process in the twentieth century.

### **5.5. Comparison between the Sixteenth and Eighteenth Centuries as well as the Nineteenth Century**

This section summarizes the findings of this chapter focusing on two issues: (1) the distribution of the variants throughout the sixteenth-eighteenth and the nineteenth centuries, and (2) the factors that favoured the incursion of the new variants. This proposes to answer the question: what has happened in the Portuguese future temporal sector from the sixteenth to the nineteenth century?

During four hundred years, the future temporal reference sector has undergone some visible changes in its representation in Portuguese. Some variants have started to recede, while others have accessed the variable context and established themselves. The synthetic future, which occurred 63% of the time in the sixteenth century, is still the most

selected variant in the nineteenth century; although, its occurrence has fallen to 53%. This same path is followed by the periphrastic variant *haver* + infinitive. Its occurrence decreased from 30% to 20% in the language.

The reverse tendency is observed among the new variants: the present and the periphrastic *ir* + infinitive. The present variant has increased its share in the variable context; moreover, in the nineteenth century it was established as the choice for the future of the verb 'to go', for which reason it was excluded from the rest of the analysis. The periphrastic *ir* + infinitive, which was taken out from the variable rule analyses of the sixteenth and eighteenth centuries due to its scarce number of occurrence (only 1%), was already preferred over the present in the future sector during the nineteenth century. As a whole, what happened in the Portuguese language during these centuries is that one simple and one periphrastic form retreated, while another simple and another periphrastic form advanced in the future sector.

Since new variants have launched into the future sector, the conditioning factors that promoted their incursion are of main interest. For that, Table 5.6 depicts a comparison between the centuries examined here. The results presented in this table were taken from Table 5.4 and 5.5. Each variant was run against all others in the multivariate analyses. Therefore, the statistical effect of each factor may change from one century to the other as the future sector is configured by different variants. Nevertheless, we detect some consistent patterns that cannot be due to chance in a period of four hundred years.

First, Table 5.6 shows that the synthetic future was originally favoured in contingent contexts, affirmative statements and with stative verbs in the sixteenth and eighteenth centuries. Then, in the nineteenth century, when all four variants were expressed

in the future sector, the synthetic future continued to be favoured by stative, and started to be preferred in two additional contexts: (i) with distal future and (ii) with non-specific adverbs. Two of its previous favouring contexts were lost in the nineteenth century: (i) *contingent* events/actions were selected as statistically significant to the occurrence of the present variant; and (ii) *affirmative* context was statistically significant to the selection of *ir* + infinitive. In the nineteenth century, both 'new' variants, i.e. the present and the periphrastic *ir* + infinitive variants, already established themselves in linguistic contexts that were once occupied by the synthetic future.

**Table 5.6:** *Seven independent variable rule analyses of the contribution of factors selected as significant to the probability that the synthetic future, the periphrastic haver + infinitive, the present and the periphrastic ir + infinitive will be selected in the future temporal context in Portuguese<sup>30</sup>.*

		Synthetic future		Haver + infinitive		Present		Ir + infinitive	
Centuries		16 <sup>th</sup>	19 <sup>th</sup>	16 <sup>th</sup>	19 <sup>th</sup>	16 <sup>th</sup>	19 <sup>th</sup>	16 <sup>th</sup>	19 <sup>th</sup>
		18 <sup>th</sup>		18 <sup>th</sup>		18 <sup>th</sup>		18 <sup>th</sup>	
Corrected mean:		.622	.545	.329	.211	.049	.098		.146
Total N:		367	268	194	104	29	48	* 31	72
Contingency	Contingent	.67		.31	.27		.83		
	Assumed	.48		.52	.52		.47		
	Range	19		21	25		36		
Sentence Type	Statement	.53		.47					
	Question	.36		.66					
	Range	17		19					
Stativity	Motion	.36	.31			.87	.89		
	Stative	.54	.52			.38	.44		
	Range	18	21			49	45		
Polarity	Affirmative	.51							.57
	Negative	.37							.09
	Range	14							48
Temp. Distance	Proximal		.37		.40	.72			.79
	Distal		.56		.55	.35			.35
	Range		19		15	37			44
Grammatical Person/Animacy	1 <sup>st</sup> Animate					.68			
	2 <sup>nd</sup> Animate					.20			
	3 <sup>rd</sup> Animate					.41			
	3 <sup>rd</sup> Inanimate					.62			
	Range					48			
Adv. Specification	No adverbial		.47				.45		.62
	Specific		.41				.88		.28
	Non-specific		.69				.46		.15
	Range		28				43		47

<sup>30</sup> In order to facilitate the visualization of the results by century, the nineteenth-century favouring factors are shaded.

<sup>31</sup> This symbol means that the variant occurred in the corpus; however, it represented only 1% of the data.

*Haver* + infinitive shows a distinct behaviour. In the sixteenth and eighteenth centuries, this variant was favoured by assumed context and questions, as opposed to the synthetic future. In the nineteenth century, assumed contexts are still statistically significant to its selection in the future sector; however, the type of sentence does not play a role anymore. At this time, distal context was selected as statistically significant to its realization, opposing it to *ir* + infinitive.

This opposition to *ir* + infinitive lead prompted the investigation of the contexts in which *haver* + infinitive had higher probability of occurrence in the nineteenth century, albeit they did not promote its choice, i.e. they were not selected as statistically significant to the occurrence of *haver* + infinitive in the future sector. The goal of this investigation was to determine whether this periphrasis had higher probability of being selected in the same contexts that *ir* + infinitive was being established in the language.

What I noticed was that *ir* + infinitive did not directly replace *haver* + infinitive; however, in the plays, *haver* + infinitive had a higher percentage of occurrence with, for instance, *stative verbs*, *affirmative contexts*, *subordinate clauses* and *no-adverbial specification*, as well as *ir* + infinitive. Thus, both periphrases showed a higher percentage of occurrences in the same linguistic contexts. Of all these factors, the loss of no-adverbial contexts seems to represent an impact on the choice of *haver* + infinitive. This factor group is particularly interesting because it demonstrates a clear, statistically significant pattern to the selection of the variants arising in the adverbial specification factor group in the nineteenth century (cf. Table 5.6). According to this pattern, each variant is preferably selected in one adverbial context. Moreover, this is the only factor group in which each of three variants has its own favoured context. It appears that the absence of *haver* + infinitive

from this pattern is evidence that it would not last in the future temporal sector in Portuguese.

Turning now to the results for the present variant, Table 5.6 shows that it is consistently favoured by verbs of motion throughout the centuries; regardless of the exclusion of the main verb 'to go' from the quantitative analyses in the nineteenth century. This is not the case for the proximal, first person and third inanimate person factors when the occurrence of this variant is examined in the nineteenth century. Based on this, it seems that this variant expanded its dominion from these primary contexts. As shown in Table 5.6, two new factors are statistically significant to its selection in the nineteenth century: (i) specific adverbs, which promote the occurrence of the present variant in the temporal adverbial configuration; and (ii) contingent events/actions, which already show that the present variant is displacing the synthetic future from its earlier preferred contexts.

Now, let us consider the last variant to enter the future temporal sector: *ir* + infinitive. As pointed out in section 5.4.4, in the nineteenth century, this variant was favoured in three linguistic contexts: affirmative utterances, proximal future eventuality and no adverbial specification. The factor that contributed the most to its selection was affirmative utterances. Recall that earlier in this chapter (cf. section 5.3) it was noted that in the sixteenth and eighteenth centuries, in which *ir* + infinitive represented only 1% of the data, this variant occurred only in affirmative contexts. The results for the nineteenth century show that *ir* + infinitive expanded its dominion through affirmative utterances and it is still avoided with negation. As noted in Table 5.6, in this process *ir* + infinitive started to be chosen in affirmative contexts, which in the sixteenth and eighteenth centuries

favoured the choice of the synthetic future in the future sector. Thus, the periphrastic *ir* + infinitive variant displaced the synthetic future from a preferred context.

The other conditioning factors reveal: (i) that the occurrence of *ir* + infinitive is favoured by absence of adverbial specification. The influence of the adverbial specification factor group was already discussed when the results of *haver* + infinitive were presented. Following the reasoning presented there, this result corroborates that *ir* + infinitive already established itself into the future sector; and (ii) that *ir* + infinitive entered the future sector by proximal future contexts, which is here defined as actions or events occurring in the same day of the utterance.

In conclusion, what we have seen in this chapter is that throughout the centuries the variants have been selected in the future context according to specific linguistic factors. The quantitative analyses effectively demonstrated the factors that contributed to their selection. They also showed how in the nineteenth century the present and the periphrastic *ir* + infinitive variants started to be preferred in some of the linguistic contexts that earlier were occupied by the synthetic future or the periphrastic *haver* + infinitive variants.

The distribution of the variants already suggests that in the twentieth century, the synthetic future, which has already lost some favoured contexts, will continue to recede (cf. Table 5.2). As we will see in the next chapter, it will be displaced from a significant context in the future sector in the twentieth century.

## Chapter 6 *Results for the Twentieth Century*

---

*(...) A vida não me chegava pelos jornais nem pelos livros*

*vinha pela boca do povo, na língua errada do povo*

*Língua certa do povo*

*Porque ele é que fala gostoso o português do Brasil (...)*

*(Fragment from *Evocação do Recife* by Manuel Bandeira 1925.)*

*'(...) Life didn't reach me through newspapers and books,  
it reached me through people's mouths, through people's wrong language  
People's right language  
Because they are the ones who speak Brazilian Portuguese so delightfully (...)*

The preceding chapter presented variable rule analyses of the use of different forms of future temporal reference in the Portuguese language in the sixteenth and eighteenth centuries as well as the nineteenth century. This chapter focuses on the variation in the future temporal sector of the written and spoken Brazilian Portuguese in the twentieth century. Its purposes are the following: (i) to determine the factors that constrain the variants in contemporary language and to compare them to the conditioning factors of the previous centuries; (ii) to examine what, if any, prescriptive influence there has been on the usage of the variants; and (iii) to assess the extent to which written representations of speech in fact are capable of portraying the spoken language. The chapter will include: (1) a presentation of the variable rule analyses of the combined written and spoken data; (2) an independent consideration of the written and the spoken data followed by a comparison of

their results; (3) a consideration of the trajectory of *ir* + infinitive; and (4) an investigation into the effect of *medium* and lastly, the effect of the level of education and of prescriptive orientation on the usage of the variants.

### 6.1. Variable Rule Analyses of the Twentieth Century Data

The twentieth century data comprise both written and spoken language. The written data were gathered from plays dated from the second half of the century onwards. The spoken data were collected from thirty-four informants from Brasilia, the capital of the country, who ranged in age from ten to thirty years old and whose level of education ranged from primary to university (cf. chapter 3 for details).

Variants of the future temporal reference are as unequally distributed in this century as they were in previous centuries. However, it is now clear that there was a reversal in the rates of usage of the variants expressing the future in Portuguese. The synthetic variant and the *haver* + infinitive periphrasis have virtually disappeared from the vernacular, while the ‘new’ variants have advanced in the variable context, taking over the role of future temporal reference. As Table 6.1 demonstrates, the present variant and the periphrasis *ir* + infinitive represent 96% of future contexts in the twentieth century data, with *ir* + infinitive being selected 85% of the time in the spoken language. Only in the written data are there more than sporadic occurrences of the synthetic variant in contemporary Brazilian Portuguese.

**Table 6.1: Distribution of the variants of the Future Temporal Reference in Brazilian Portuguese by written and spoken data from the twentieth century.**

	<i>Synthetic future</i>		<i>Haver + infinitive</i>		<i>Present</i>		<i>Ir + infinitive</i>		Total
	N	%	N	%	N	%	N	%	
<b>Written</b>	46	9	5	1	93	18	384	73	528
<b>Spoken</b>	4	1	—	—	104	14	613	85	721
<b>Total</b>	50	4	5	0	197	16	997	80	1249

The difference in the distribution of the variants in the written and the spoken data suggests that *ir + infinitive* has advanced further in the spoken language than in the written. This inference will be empirically tested by the comparison of the results of the variable rule analyses in the following sections of this chapter.

#### 6.1.1. Factors

The variable rule analyses of the twentieth century data replicates the linguistic factors that were tested in the previous centuries. A single addition was made, however, to test whether or not medium, either spoken or written, constrains the choice of variants (cf. the working hypotheses in chapter 4). As well, level of education was considered as a factor in the spoken data, which will be analyzed in another section.

#### 6.1.2. Results

Three variable rule analyses were run in order to verify the influence of the postulated linguistic factors in the choice of the three remaining variants: the synthetic, the

present and *ir* + infinitive variants in the future variable<sup>32</sup> context in contemporary Portuguese.

As noted in the previous chapter, the future of the main verb *ir* ‘to go’ was rarely expressed by the *ir* + infinitive periphrasis (cf. Table 4.4.). There were only two such occurrences and they were in the speech of a ten-year-old informant. With the decline of the synthetic future, the present variant has enlarged its context to include future constructions with *ir* ‘to go’ as main verb. As a result, in the twentieth century data, 97% of the time that the main verb *ir* ‘to go’ was used in the future temporal context, it was expressed by the present variant.

Table 6.2 shows the distribution of the variants according to verbs class, either stative, “to go”, or other verb of motion. Because of the near categorical behaviour of the verb *ir* ‘to go’, it was decided to exclude it from the variable rule analyses of the twentieth century data.

**Table 6.2: Distribution of the variants of the Future Temporal Reference in Brazilian Portuguese according to stativity in the twentieth century.**

	<i>Synthetic future</i>		<i>Present</i>		<i>Ir + infinitive</i>		Total
	N	%	N	%	N	%	
<b>Stative</b>	45	4	90	9	893	86	1028
<b>Other Motion</b>	4	3	29	21	102	76	135
<b>Ir ‘to go’</b>	1	1	78	97	2	2	81
<b>Total</b>	50	4	197	16	997	80	1244

<sup>32</sup> The periphrastic *haver* + infinitive variant was disregarded in the quantitative analyses of this century due to its scarce number of occurrences.

Table 6.3 depicts the results of the variable rule analyses for all tokens of the twentieth century, excluding the main verb 'to go'. The most significant factor group affecting the selection of the variants is *medium*. As Table 6.3 shows, in the twentieth century the selection of the synthetic future is highly favoured by the written data. Notice that this factor contributes the most to its choice, as expressed by a range of 58. The effect of this group of factors on the realization of the periphrastic variant *ir* + infinitive is also the most significant (range of 36); however, it affects this variant on the opposite direction, i.e. *ir* + infinitive is favoured in speech.

The present variant is also favoured in the written data. Thus, *medium* also exerts influence on its realization; however, the range of this group of factors (range of 25) is smaller than that of *adverbial specification* (range of 34) and *contingency* (range of 30), which demonstrates that *medium* is not the most important constraint influencing the selection of the present variant in twentieth-century Brazilian Portuguese.

Considering the results of *medium* for the synthetic future and the periphrasis, it was decided to examine the written and the spoken data separately. The next sections show the analyses.

**Table 6.3: Three variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, the Present and the periphrastic variant *Ir + infinitive* will appear in future temporal reference contexts in the twentieth century in Brazilian Portuguese. Without the main verb *ir* 'to go'.**

	<i>Synthetic Future</i>		<i>Present</i>		<i>Ir + infinitive</i>	
Total N (1163)	49		119		995	
Corrected Mean	.018		.102		.856	
	% N		% N		% N	
<b>Medium</b>						
Writing	.82	9 45	.64	14 72	.30	77 384
Speech	.24	1 04	.39	7 47	.66	92 611
Range	58		25		36	
<b>Grammatical Person/Animacy</b>						
1 <sup>st</sup> Animate	.46	3 13			.46	86 397
2 <sup>nd</sup> Animate	.37	2 03			.64	92 144
3 <sup>rd</sup> Animate	.46	3 11			.56	88 284
3 <sup>rd</sup> Inanimate	.71	10 21			.38	77 170
Range	34				26	
<b>Adverbial Specification</b>						
No adverbial			.43	7 65	.57	89 770
Specific			.77	23 33	.27	75 107
Non-specific			.60	14 21	.36	77 118
Range			34		30	
<b>Contingency</b>						
Assumed			.47	9 95	.53	87 924
Contingent			.77	24 24	.21	70 71
Range			30		32	
<b>Stativity</b>						
Stative			.47	9 90	.52	87 893
Motion			.71	21 29	.32	76 102
Range			24		20	
<b>Type of Clause</b>						
Main			.54	12 119	.47	84 765
Subordinate			.37	6 14	.61	91 230
Range			17		14	
Factor groups not selected as significant:						
Grammatical Person/Animacy				X		
Adverbial Specification	X					
Contingency	X					
Stativity	X					
Type of Clause	X					
Temporal Distance	X		X		X	
Polarity	X		X		X	
Sentence Type	X		X		X	
Presence of Clitics	X		X		X	

## 6.2. Written Data Analyses

The quantitative analyses of the written data demonstrate that a total of five linguistic groups of factors influence the choice of the variants in the future context in Brazilian Portuguese. Table 6.4 displays the results of these analyses. The following discussion of the results will show which variant is favoured by each of these factors.

### 6.2.1. *The Synthetic Future*

The first variant to be so analyzed is the synthetic future. Recall that the overall rate of this variant in the plays is only 9%, which shows that the usage of this variant has declined in the language; however, more revealing than the overall rate is the analysis of the factors conditioning its usage in the twentieth century. As shown in Table 6.4, the synthetic future is constrained by two groups of factors, namely *stativity* and *grammatical person/animacy*.

*Stativity* is the most significant group of factor on the choice of the synthetic future, with a range of 36. Table 6.4 indicates that *stative verbs* favour this variant, while verbs of motion strongly disfavour it. Recall that this factor group has been selected as constraining the usage of the synthetic future since the sixteenth and eighteenth century analysis, and it has always affected this variant in the same direction, i.e. the synthetic future has always been favoured by stative verbs. The magnitude of its effect, which is measured by the range, however, has increased throughout the centuries. I suggest that the consistency and the increase of the effect of this group of factor indicate that there is a strong co-relation between the stativity of the verb and the choice of the synthetic variant for future expression in Portuguese.

**Table 6.4: Three variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, the Present and the periphrastic variant *Ir + infinitive* will appear in future temporal reference contexts in the twentieth century in the written data. Without the main verb *ir 'to go'*.**

	<i>Synthetic Future</i>	<i>Present</i>	<i>Ir + infinitive</i>
Total N (501)	45	72	384
Corrected Mean	.090	.144	.766
	% N	% N	% N
<b>Stativity</b>			
Stative	.54 10 44	.46 11 52	.52 79 358
Motion	.18 02 01	.83 43 20	.28 55 26
Range	36	37	24
<b>Grammatical Person/ Animacy</b>			
1 <sup>st</sup> Animate	.47 08 13	.68 21 36	.39 71 123
2 <sup>nd</sup> Animate	.38 05 03	.12 02 01	.81 93 52
3 <sup>rd</sup> Animate	.44 06 10	.47 12 19	.55 81 125
3 <sup>rd</sup> Inanimate	.68 16 19	.50 13 16	.43 71 84
Range	30	49	42
<b>Adverbial Specification</b>			
No adverbial		.44 11 40	.56 81 306
Specific		.77 36 16	.31 60 27
Non-specific		.61 20 16	.34 64 51
Range		33	25
<b>Type of Clause</b>			
Main		.54 16 65	
Subordinate		.33 07 07	
Range		21	
<b>Contingency</b>			
Assumed			.52 78 364
Contingent			.27 59 20
Range			25
<b>Factor groups not selected as significant:</b>			
Grammatical Person/Animacy		X	
Adverbial Specification	X		
Contingency	X		
Stativity	X		
Type of Clause	X		
Temporal Distance	X	X	X
Polarity	X	X	X
Sentence Type	X	X	X
Presence of Clitics	X	X	X

*Grammatical person/animacy* also contributes to the selection of the synthetic future. As we can see in Table 6.4, *3<sup>rd</sup> person inanimate subjects* favour the choice of this variant. Although this constraint was not previously found to be significant in the selection of the future variant, it was relevant to the selection of the present variant in the sixteenth and eighteenth centuries. Now, in the twentieth-century plays, a favouring effect of inanimate subjects emerges for the synthetic future, precisely as this variant is receding from other areas of the future sector. It seems that this factor functions as a last resort for the synthetic future, a context where neither of the competing variants, *ir* + infinitive or the present variant, is particularly favoured. It is difficult, however, to provide an explanation for this affinity, as grammatical person/animacy has not shown a regular effect on the variants.

In summary, the synthetic future is retained as variant of the future temporal reference in Portuguese favoured in contexts with *3<sup>rd</sup> person inanimate subjects* and with stative verbs.

Having uncovered the historical and contemporary factors significant to the selection of the synthetic future, it is now possible to compare those results to the prescriptive tradition that describes the use of this variant. First, it is clearly evident that the usage of the synthetic future runs counter to the grammarians' widespread observations that it is the default form for expressing the future temporal reference in Brazilian Portuguese. Secondly, the list of meanings associated to this variant in the twentieth-century grammars shows no relation with the empirical findings. As depicted in Table 2.2 (cf. chapter 2), starting from nearly the second half of the twentieth century, grammarians ascribed various meanings to the synthetic future, namely *uncertainty*, *doubt*, *probability*, *possibility*,

*politeness, suggestion, intention, determination, certainty* and *promise*. These numerous readings do not reflect the linguistic factors that motivate the selection of the synthetic future (cf. Table 6.4). Thus, no possible relation can be established between prescriptive orientation and the usage of the synthetic future in written Brazilian Portuguese.

### 6.2.2. *The Present*

The second variant to be analyzed is the present. Table 6.4 shows that in the twentieth century four factors are statistically significant to the choice of this variant in the written data: *grammatical person/animacy, stativity, adverbial specification* and *type of clause*.

The factor group that contributes the most to the selection of the present variant in the future sector is *grammatical person/animacy*, as expressed by the range of 49. *1<sup>st</sup> person animate subjects* favour the occurrence of this variant with the probability of .68. Recall that in the sixteenth and eighteenth centuries, this group of factor was also selected as statistically significant to the occurrence of this variant. One possible explanation for this affinity is interaction between factor groups. Looking more closely at the twentieth century written data, it was determined that verbs of motion, which strongly favour the present variant, show a high likelihood of co-occurring with first person subjects. This association is likely the cause of the selection of the 1<sup>st</sup> person subjects as favouring the present variant. However, the cross-tabulation did not show empty cells, i.e. total interaction between two factors.

*Stativity* also exhibits a significant influence on the selection of the present variant. Note that *verbs of motion* highly favour the occurrence of this variant in the future sector. Recall that in the twentieth century, the main verb *ir* 'to go' selected the present variant in

future contexts (cf. Table 6.2). Because of this near categorical behaviour, it was decided to exclude *ir* 'to go' from the quantitative analyses. Even after this exclusion, the present variant is still favoured by verbs of motion. The most frequent verbs of motion in the data, for example, *voltar* 'to come back', *vir* 'to come' and *passar* 'to pass by', favour the selection of the present variant<sup>33</sup>, although they also co-occur with *ir* + infinitive.

The next factor group selected as statistically significant to the choice of the present variant is *adverbial specification*. As we can see in Table 6.4, this variant is favoured by both *specific* and *non-specific* adverbials, with the probability of .77 and .61, respectively. The association of adverbial specification with the present variant has already been explained in this study as a result of a disambiguating effect. Adverbials confer to the context the future temporal reference reading; thus, the future eventuality is not interpreted as a present one.

The last group of factor selected by the quantitative analysis is *type of clause*. As shown in Table 6.4, *main clauses* favour the choice of the present variant in future contexts. This group of factors was found to be statistically significant to the occurrence of this variant only in the written data of the twentieth century. These results were compared to Comrie's (1985) claim that the present tense expresses future time in main clauses in English when there is a scheduled situation, *in verbis*:

---

<sup>33</sup> The cross-tabulation of stativity and adverbial specification did not show interaction. It tested the hypothesis that all verbs of motion co-occurred with adverbial specification due to scheduled events.

In main clauses, there is a heavy constraint on the use of the present tense with future time reference, namely that the situation referred to must be one that is scheduled. In our example above, *the train departs at five o'clock tomorrow morning*, use of the present tense is justified because the situation referred to is indeed one that is scheduled. (Comrie 1985: 47)

This explanation of the influence of main clauses on the selection of the present is related to adverbial specification. If we interpret Comrie's concept of scheduled events/actions as those that express a specific future time reference by means of an adverbial specification, one can assume that adverbial specification will also favour the use of the present variant in Portuguese. And indeed, specific adverbs do favour the selection of the present. In order to ensure that these two factors represented independent effects and were not the result of a tendency for adverbial specification to occur in main clauses, results for clause type and adverbial specification were cross tabulated. The cross-tabulation did not show interaction<sup>34</sup>, which means that in Brazilian Portuguese, the present variant co-occurs with adverbial and no-adverbial specification in both main and subordinate clauses. Moreover, its usage is highly constrained by specific adverbs in both types of clauses. Thus, the selection of main clauses as statistically significant to the choice of the present variant is not an epiphenomenon of the effect of adverbial specification.

In order to find a reason for the selection of type of clause as statistically significant to the choice of the present variant only in the written data of the twentieth century, it was

---

<sup>34</sup> The tokens with the present variant in main clauses are distributed as follow: 37 tokens with no adverbial specification (12%), 15 tokens with non-specific adverbs (23%), and 13 tokens with specific adverbs (35%). In subordinate clauses, they are distributed as follows: 3 tokens with no adverbial specification (4%), 1 token with non-specific adverbs (7%), and 3 tokens with specific adverbs (38%).

decided to examine the relative weights of all variants in main and subordinate clauses. I noticed that the periphrastic variants consistently showed higher relative weights in subordinate clauses, while the non-periphrastic variants went in the opposite direction with higher relative weights in main clauses. There were two exceptions: (i) the present in the sixteenth and eighteenth centuries, when it apparently began its incursion in the variable context with a very low frequency (5%); and (ii) the synthetic future in the written data of the twentieth century, when it recedes. Observe Table 6.5.:

**Table 6.5: Relative weights of the variants in main and subordinate factors in the written data from the sixteenth to the twentieth centuries. (Relative weights in brackets mean that the factors were not selected as statistically significant)**

Centuries	Synthetic Future			Haver + infinitive			Present			Ir + infinitive		
	16 <sup>th</sup> 18 <sup>th</sup>	19 <sup>th</sup>	20 <sup>th</sup> written	16 <sup>th</sup> 18 <sup>th</sup>	19 <sup>th</sup>	20 <sup>th</sup> written	16 <sup>th</sup> 18 <sup>th</sup>	19 <sup>th</sup>	20 <sup>th</sup> written	16 <sup>th</sup> 18 <sup>th</sup>	19 <sup>th</sup>	20 <sup>th</sup> written
Corrected Mean	.622	.545	.090	.329	.211		.049	.098	.144		.146	.766
Total N	367	268	45	194	104		29	48	72		72	384
<i>Type of Clause</i>												
Main	[.51]	[.51]	[.50]	[.49]	[.48]		[.45]	[.54]	.54		[.47]	[.48]
Subordinate	[.45]	[.42]	[.50]	[.52]	[.59]		[.66]	[.27]	.33		[.64]	[.59]
Range									21			

Note that the ranges, i.e. the differences between the relative weights of main and subordinate factors, are always higher for the present variant. Thus, the present variant is the most affected variant by type of clause, although it is statistically significant to its choice only in the twentieth century.

As can be seen in Table 6.5, the synthetic future has higher relative weights in main clauses in the sixteenth and eighteenth century as well as in the nineteenth century. In the

twentieth century, however, there is no difference in the relative weights (.50), which means that *type of clause* is totally neutral with regard to the selection of the synthetic variant in the future sector. Thus, in main clauses the most available option for representing the future is the present variant. Therefore, the effect of main clauses on the present variant becomes significant. The options offered by the linguistic system to represent the future explain the preferred selection of the present variant in main clauses in the twentieth century<sup>35</sup>.

Now it is possible to compare these results with those of the previous centuries. As already discussed, grammatical person/animacy has not shown a consistent effect on the present variant, although it was also significant in the sixteenth and eighteenth centuries. Another factor which emerges only in the twentieth century was the effect of type of clause. Some explanatory possibilities were examined for such a result, and it was suggested that the selection of main clause as statistically significant to the present variant is related to the synthetic future's withdrawal from the future sector.

Stativity and adverbial specification have shown a consistent effect on the realization of the present variant. Verbs of motion have favoured the present variant since the sixteenth and eighteenth centuries when it first appeared in the future temporal context in this study. However, the importance of this group of factors has lessened throughout the centuries, as indicated by its declining range: from 49 in the sixteenth and eighteenth centuries, to 45 in the nineteenth century and 37 in the written twentieth century.

---

<sup>35</sup> I examined the contribution of *contingency* to the occurrence of the present variant. This group of factor was selected as statistically significant to the choice of this variant in the nineteenth century, but not in the twentieth century. However, both the present variant and the synthetic future have a higher probability of occurrence in contingent contexts. It is clear that they compete for this portion of the future temporal reference sector. Furthermore, what emerges from the relative weights is that these variants compete with each other for a restricted context in the language, while the periphrastic variant *ir* + infinitive advances in a more recurring future context, i.e. assumed events/actions.

The favouring effect of adverbial specification on the occurrence of the present variant was reported in the nineteenth century. Recall that in the nineteenth century specific adverbs were already statistically significant to the selection of the present variant. In the twentieth century, however, the present extended its dominion over the contexts with non-specific adverbs. Comparing the results obtained from the other centuries with those of the twentieth century, it is evident that the present variant occupies the adverbial context that was previously selected as significant in the choice of the synthetic form, while the synthetic usage retreats in the Portuguese language. Based on the trajectory of the variants throughout the centuries, the prediction is that if the variation in the future sector does not remain stable and the observed changes reach completion, there will be only the opposition of no-adverbial and adverbial specification, with the periphrastic variant *ir* + infinitive continuing to encroach in future contexts with adverbial specification.

An interesting finding of this study is that in the twentieth century the effect of adverbial specification on the present variant in Brazilian Portuguese plays is identical to that reported to spoken Canadian French by Poplack and Turpin (1999: 149). Both languages share the same effect of adverbial specification, and most importantly, they share the same ranking of factors in this group. In French, the present variant is also favoured by specific and non-specific adverbs. This finding suggests that this relationship may be also found in other Romance languages.

The last consideration to be made in this section is whether the grammarians' remarks on the usage of the present variant reflect its usage in Brazilian Portuguese. As summarized in Table 2.2 (cf. chapter 2), in the twentieth century numerous meanings are attributed to the present variant in Portuguese. *Certainty, necessary occurrence, emphasis,*

*firm intention, promise, high probability, unfeasibility, impossibility, politeness* and *proximity* are all meanings presented as explaining the usage of this variant in the future sector. As we can see in Table 6.4, none of these readings are in fact related to the selection of the present variant. Although some attitudes as certainty and firm intention are commonly expressed in the 1<sup>st</sup> person, we saw that the selection of the 1<sup>st</sup> person is better explained by its co-occurrence with the usage of verbs of motion in the data. The only exceptions are remarks made by Cunha & Cintra (1984/ 1999: 438), Thomas (1969: 116) and Van Achter et al. (1996: 155). These books, from which two are addressed to non-native speakers (cf. Table 2.1), clearly prescribe that the usage of the present form requires temporal adverbs in order to have the action clearly situated in the future. The empirical analysis corroborates this claim since adverbial specification contributes to the choice of the present variant in the future context.

### 6.2.3. *The Periphrastic Variant Ir + infinitive*

This section addresses the most recent variant to extend its dominion into the future sector in Brazilian Portuguese. By and large, the periphrastic variant *ir + infinitive* represents the preferred choice to express future temporal reference in the written data of the twentieth century (cf. Table 6.1). Five factors highly contribute to its selection over the present variant and the synthetic future: 2<sup>nd</sup> and 3<sup>rd</sup> animate persons, no adverbial specification, assumed events/actions and stative verbs.

The *grammatical person/animacy* factor group highly constrains *ir + infinitive*. The factors of 2<sup>nd</sup> and 3<sup>rd</sup> animate subjects favour the selection of this periphrasis, with the probability of .88 and .55, respectively. The disfavoured effect of inanimate subjects, however, is surprising. It was not expected that this factor would not have any effect at all

on the periphrasis in the twentieth century. This statistical effect likely reflects the retention of the synthetic future in this context.

*Adverbial specification* is also selected as statistically significant to the realization of this periphrasis (range of 25). Table 6.4 shows that *ir* + infinitive is favoured by *no adverbial specification*. As we can see, in the twentieth century, preference for *ir* + infinitive and the present variant differ according to type of adverbial modification.

The next group of factor to influence the choice of this periphrasis is *contingency*. Recall that this group is composed by two factors: contingent events/actions and *assumed* ones. The latter favours the occurrence of *ir* + infinitive in the future sector. Thus, in Brazilian Portuguese this periphrasis is preferred when the future eventuality is not dependent upon any circumstances, i.e. it is less hypothetical.

The last significant group of factor is *stativity*. As can be seen in Table 6.4, *stative verbs* favour *ir* + infinitive with the probability of .52. In the grammaticalization process of this periphrasis as a future marker, the usage of the auxiliary *ir* 'to go' with another verb of motion implies that the lexical content of the auxiliary is lost. Thus, when such an association happens the periphrasis is considered to be advanced in the process of grammaticalization. As we can see, that is still not the case in the twentieth century written data. The analysis shows that despite the high frequency of *ir* + infinitive in future contexts, there is still a propensity to avoid this periphrasis with verbs of motion. Therefore, its process of grammaticalization was not well-advanced in the language at that time.

How are these results related to those of the nineteenth century written data? First recall that in the nineteenth century *ir* + infinitive entered the future sector. At that time, it was favoured by *affirmative utterances*, *no adverbial specification* and *proximal future*

*eventuality* (cf. Table 5.5). In the twentieth century written data, this periphrasis is already the preferred choice to express the future temporal reference. It was favoured by 2<sup>nd</sup> and 3<sup>rd</sup> animate subjects, no adverbial specification, assumed events/actions and stative verbs (cf. Table 6.4). As we see, no adverbial specification has constantly favoured this variant. Affirmative utterances and proximal future eventuality were also primary contexts for the realization of this periphrasis. As suggested in chapter 5, *ir* + infinitive next extended its usage to negative utterances and distal contexts. Finally during the process of expansion in the future sector, in the twentieth century it occupied the context of assumed events/actions, that earlier had favoured the periphrastic *haver* + infinitive variant (cf. Table 5.5). As the synthetic future recedes, *ir* + infinitive started to be favoured by stative verbs, that earlier favoured only the synthetic future (cf. Table 5.5). Only the effect of grammatical person/animacy is in fact unclear when we examine the trajectory of *ir* + infinitive from the nineteenth to the twentieth century in the written data.

The significant effect of no adverbial specification on the choice of *ir* + infinitive parallels the result found by Poplack and Turpin (1999: 149). According to these authors, this periphrasis is favoured by no adverbial specification in Canadian French as well. As already mentioned in the previous section, in twentieth-century written Brazilian Portuguese and spoken Canadian French, *ir* + infinitive and the present variants are favoured by the same factors in the type of adverbial specification group of factors.

The results reported here cannot be related to the meanings associated to this periphrasis by grammarians (cf. Table 2.2). From the second half of the twentieth century onwards, grammarians started to acknowledge the usage of the periphrastic *ir* + infinitive variant in Portuguese; however, they assigned this variant the meanings of *movement*,

*proximity, intention, resolution, certainty* and *necessary occurrence*. Recall that every time *ir* + infinitive meant physical movement, it was not considered a variant of future time reference; accordingly, it was not included in the quantitative analyses of this study (cf. section 4.2). Proximity, frequently evoked in the literature, was not confirmed by the analysis, i.e. temporal distance does not constrain the usage of this variant in the language. The remaining meanings, i.e. intention, resolution, certainty and necessary occurrence, could be interpreted as reflecting the preferred usage of this variant in assumed events/actions; however, in such a case, these meanings should not be attached to the present variant, which is not favoured by assumed events/actions. Note that the grammarians confer them to both variants (cf. Table 2.2). Therefore, I conclude that the prescriptive orientation does not reflect the usage of *ir* + infinitive.

The next section will present a summary of the trajectories of the variants from the sixteenth to the twentieth centuries in Portuguese.

### 6.3. Comparison of the Results of the Written Data: 16<sup>th</sup> – 19<sup>th</sup> Centuries vs. 20<sup>th</sup> Century

The first noticeable finding of the analyses of the twentieth century's written data is the dramatic change in the frequency of the synthetic future and the periphrastic variant *ir* + infinitive. According to the written data, there was a reversal in the rates of usage of these two variants in the language from the beginning of the nineteenth century to the end of the twentieth century. Moreover, *haver* + infinitive was definitely excluded from the variable context in the twentieth century (only 1% of the data). Table 6.6 depicts the selected factor groups in the written data throughout the centuries.

**Table 6.6: Ten independent variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, the periphrastic variant Haver + infinitive, the Present and the periphrastic variant Ir + infinitive will be selected in the future temporal context in Portuguese.**

	Synthetic future			Haver + infinitive			Present			Ir + infinitive		
	16 <sup>th</sup> 18 <sup>th</sup>	19 <sup>th</sup>	20 <sup>th</sup>	16/18 <sup>th</sup>	19 <sup>th</sup>	20 <sup>th</sup>	16 <sup>th</sup> 18 <sup>th</sup>	19 <sup>th</sup>	20 <sup>th</sup>	16 <sup>th</sup> 18 <sup>th</sup>	19 <sup>th</sup>	20 <sup>th</sup>
Centuries												
Corrected mean:	.622	.545	.090	.329	.211		.049	.098	.144		.146	.766
Total N:	367	268	45	194	104		29	48	72	*	72	384
Contingency	.67			.31	.27			.83				.27
Contingent Assumed	.48			.52	.52			.47				.52
Range	19			21	25			36				25
Sentence Type	.53			.47								
Statement Question	.36			.66								
Range	17			19								
Stativity	.36	.31	.18				.87	.89	.83			.28
Motion Stative	.54	.52	.54				.38	.44	.46			.52
Range	18	21	36				49	45	37			24
Polarity	.51										.57	
Affirmative Negative	.37										.09	
Range	14										48	
Temp. Distance		.37			.40		.72					
Proximal Distal		.56			.55		.35				.79	
Range		19			15		37				44	
Grammatical Person/Animacy			.47				.68		.68			.39
1 <sup>st</sup> Animate			.38				.20		.12			.81
2 <sup>nd</sup> Animate			.44				.41		.47			.55
3 <sup>rd</sup> Animate			.68				.62		.50			.43
3 <sup>rd</sup> Inanimate			30				48		49			42
Range									.44			.56
Adv. Specification		.47						.45	.44		.62	
No adverbial Specific		.41						.88	.77		.28	.31
Non-specific		.69						.46	.61		.15	.34
Range		28						43	33		47	25
Type of Clause												
Main Subordinate									.54			
Range									.33			
									21			

Comparing the statistically significant factors conditioning the synthetic future in the previous centuries to the ones in the twentieth century, it is evident that this variant has receded from the following significant contexts over the years: statements, contingent events/actions, affirmative utterances and distal future. In the twentieth century, this variant also lost its preferred status with temporal adverbs. (cf. Table 6.6). Contingent event/action and non-specific adverbial contexts were occupied by the present variant, while affirmative utterances were taken on by *ir* + infinitive. In the twentieth century written data, the synthetic future is preferred only by stative verbs and 3<sup>rd</sup> person inanimate subjects.

The present variant has shown a very regular behaviour in the variable context. Its development has been marked by its tendency to express the future with verbs of motion and to be selected in restricted contexts, namely in contingent future events/actions and utterances with temporal specification. In addition, its usage has also been preferred with 1<sup>st</sup> animate person and 3<sup>rd</sup> person inanimate subjects. Since the sixteenth century, the present has steadily increased its frequency in written data. However, such an increase has been minor when compared to the increase of *ir* + infinitive.

The steady rise of the present variant appears to be limited to a few specialized contexts. The opposite is the case with the expansion of *ir* + infinitive. In general, the statistically significant factors that have determined its choice represent the most recurring contexts in the data. This observation does not elucidate what has motivated the change in the future temporal reference sector in Portuguese. However, once this variant became well-entrenched in the language, the fact that it was favoured by such recurring contexts may have accelerated its rapid increase in frequency<sup>36</sup>. If we compare its occurrences in the

---

<sup>36</sup> I am aware that social factors also have an effect on linguistic change. Likewise, variation and change in other sectors of the language might also contribute to the change in the future temporal context; for instance,

twentieth century with those of the nineteenth century, we notice that in the twentieth century it has already expanded its usage beyond the affirmative and proximal contexts in which it was previously favoured (cf. Table 6.6). Recall that in the nineteenth century both these factors were selected as statistically significant to its occurrence, with the relative weights of .57 and .79, and the ranges of 48 and 44, respectively. Already by the twentieth century, the periphrasis was no longer restricted to affirmative contexts. Moreover, its strong association to the proximal future was no longer evident.

Another interesting result of the expansion of *ir* + infinitive is that it occupied assumed contexts, which had previously favoured the periphrastic variant *haver* + infinitive in the sixteenth and eighteenth as well as in the nineteenth century. In addition, stative verbs are also statistically significant to its selection (cf. Table 6.6), which suggests that unless the ongoing change stabilizes, the synthetic future will be excluded from this context by the periphrastic *ir* + infinitive variant.

The next section will turn this investigation to contemporary spoken Brazilian Portuguese, and then compare the results with those of the twentieth century written data.

#### **6.4. Variable Rule Analysis of the Spoken Data**

This section focuses on the last variable rule analysis of this study. Differently from the previous analyses, this analysis is based on traditional resource of variationist studies: spoken language. The results here are essential to the understanding of the process of change in the future temporal sector in Portuguese.

---

the disappearance of *haver* as an auxiliary in Brazilian Portuguese is a phenomenon that seems not to be restricted to the future sector.

As we saw in Table 5.2, the exponents of future temporal reference in spoken Brazilian Portuguese are virtually reduced to the periphrasis formed by the present form of *ir* 'to go' plus the infinitive of a main verb. *Ir* + infinitive was chosen 85% of the time the future was meant by speakers. If we exclude the future occurrences with the main verb 'to go', which were all virtually expressed by the present variant (cf. Table 6.2), the overall rate of *ir* + infinitive increases to 93% of future temporal reference.

The distribution of the variants (cf. Table 6.1) showed that the synthetic future accounts for only 1% of the spoken data, reminiscent of the rate of *ir* + infinitive when it first emerged in the variable context in the sixteenth century (cf. Table 5.2). Due to the low frequency of the synthetic future in the spoken data, it was excluded from the quantitative investigation. In general terms, it can be considered to have been overridden in the future context. Consequently, in the quantitative analysis the present variant, which is productive in a very restricted context, is compared to *ir* + infinitive, the overall preference.

Given that the primary goal was to examine the factors that are statistically significant to the selection of the present variant, special attention was paid to contexts in which the present variant is maintained in contemporary spoken language. Table 6.7 displays the factors selected as statistically significant to the probability that the present variant will be chosen in the future sector in spoken Brazilian Portuguese. Two general results are first noticeable in this Table: (i) both groups of factors selected here also constrained the usage of the variants in the written data; (ii) the factors influence the selection of the variant in the same direction.

As shown in Table 6.7, contingent contexts and adverbial specification retain the present variant in the language. In other words, these two factors are the last barriers to the

expansion of *ir* + infinitive. The resistance of these particular factors to encroachment by the periphrastic form is best explained by returning to the historical trajectory of the variants, as presented in the quantitative analyses.

**Table 6.7: Variable rule analysis of the contribution of factors selected as significant to the probability that the Present variant will appear in future temporal reference contexts in contemporary spoken Brazilian Portuguese. Without the main verb *ir* 'to go'.**

	<i>Present</i>		
Total N (658)	.47		
Corrected Mean	.071		
		%	N
Contingency			
Contingent	.87	25	17
Assumed	.45	5	30
Range	42		
Adverbial Specification			
Specific	.79	18	17
Non-specific	.60	7	5
No adverbial	.42	5	25
Range	37		

**Factors not selected as significant:** Grammatical Person/Animacy, Stativity, Temporal Distance, Presence of Clitics, Polarity, Sentence Type, Type of Clause.

First, recall that the preferred variant in contingent contexts was the synthetic future. From the beginning, the present variant competed with the synthetic future for this context; thus, it has been one targeted locus of the present, which definitely assumed it as the synthetic variant receded in the language. It follows from the history of the future temporal sector in Portuguese that *ir* + infinitive has no tradition in such a context. Thus, it is no surprise it is still disfavoured in there.

Second, the present variant has been associated with adverbial specification for centuries in Portuguese. It has expanded its usage over contexts with non-specific adverbs, and after this expansion, a clear dichotomy was established in the system: the less frequent future temporal contexts have adverbial specification and are rather expressed by the present variant; while the more frequent future temporal contexts have no adverbial specification and are preferably conveyed by the periphrasis. As a result, the present variant holds on less frequent, very specific, and consequently more salient contexts, which are naturally expected to offer more resistance to the infiltration of the periphrasis. Indeed, a hallmark of these two contexts is that they have marked an opposition between these variants since their incursion in the future context. Observe Table 6.8.

**Table 6.8: The trajectory of the Present and the periphrastic variant *Ir + infinitive* with the factor groups selected as statistically significant in contemporary spoken Brazilian Portuguese (adapted from Tables 5.4, 5.5, 6.4, 6.7)**

	<i>Present</i>				<i>Ir + infinitive</i>			
<i>Centuries:</i>	16 <sup>th</sup> 18 <sup>th</sup>	19th	20th written	20th spoken	16 <sup>th</sup> 18 <sup>th</sup>	19th	20th written	20th spoken
Corrected Mean	.049	.098	.144	.071		.146	.766	.929
<i>Contingency</i>								
Contingent		.83		.87			.27	.13
Assumed		.47		.45			.52	.55
Range		36		42			25	42
<i>Adverbial Specification</i>								
Specific		.88	.77	.79		.28	.31	.21
Non-specific		.46	.61	.60		.15	.34	.40
No adverbial		.45	.44	.42		.62	.58	.58
Range		43	33	37		47	25	37

Table 6.8 depicts the relative weights of the present and *ir* + infinitive variants according to *contingency* and *adverbial specification*. Note that contingent contexts favour the selection of the present in Portuguese. The same holds to contexts with specific adverbial specification. Moreover, in the twentieth century written data, when the synthetic future is no longer productive, contexts with non-specific adverbs also favour the choice of the present variant in the future sector. *Ir* + infinitive, on the contrary, has been favoured in assumed contexts and by no-adverbial specification. Considering the consistency, I conclude that both group of factors have been deeply connected to the occurrence of the present and *ir* + infinitive variants in the language

The next section will compare the results from the spoken and written data to determine whether the two mediums represent the same stage in the development of the future temporal reference sector.

### 6.5. Comparison: Twentieth Century Written and Spoken Data

Comparison between the distribution of the variants in the written and the spoken data reveals salient differences (cf. Table 6.9). The synthetic future, which represents 9% of the occurrences in the written data, accounts for less than 1% of the future contexts in the spoken data, and consequently could not be analyzed by the quantitative analysis. This effectively reduced the number of forms of future temporal reference to two variants in speech: the present and the periphrastic form *ir* + infinitive. The rate of *ir* + infinitive is 77% in the written data, compared to 93% in the spoken one. Hence, future temporal reference is expressed almost always by *ir* + infinitive in spoken Brazilian Portuguese.

More important to this study than the overall rates, however, is the comparison of the linguistic factors that constrain the usage of each variant. This comparison enables us to pinpoint the similarities and differences in their conditioning. Regardless of rate, it is expected that the statistical significance and the direction of effect of the factors will be similar, since both the written and the spoken data represent the usage of the same language in the same century.

Table 6.9 depicts the relative influence of the linguistic factors in the written and spoken data of the twentieth century. This Table shows that the groups of factors selected as statistically significant in the written and spoken data are not exactly the same. *Stativity*, *grammatical person/animacy* and *type of clause* contribute no significant effect to variant selection in the spoken data. However, both *adverbial specification* and *contingency* were

selected as statistically significant, and showed the same direction of effect<sup>37</sup> in the spoken and written data.

As discussed previously, one possible explanation for the discrepancy between the spoken and written data is the possibility that the written data are not influenced by the same conditioning factors as speech. Another less dramatic conclusion is that the spoken language merely represents a slightly more advanced state of changes that are not yet evident in written language.

---

<sup>37</sup> Quantitative analysis of the twentieth-century written data was performed taking into account only the tokens of the present and the periphrastic *ir* + infinitive variants. The groups of factors selected as statistically significant were the same presented in Table 6.9 plus *contingency*. This finding was consistent with the finding reported to the spoken data: the present was favoured by *contingent* events/actions, while the periphrasis was favoured by *assumed* events/actions.

**Table 6.9:** *Five independent variable rule analyses of the contribution of factors selected as significant to the probability that the Synthetic Future, the Present and the periphrastic Ir + infinitive variants will be selected in future temporal contexts in Brazilian Portuguese in the twentieth century. Brackets indicate that the factor group was not selected as significant (adapted from Tables 6.5 and 6.7)*

		<i>Synthetic future</i>		<i>Present</i>		<i>Ir + infinitive</i>	
		Writing	Speech	Writing	Speech	Writing	Speech
Corrected mean:		.090	*	.144	.071	.766	.929
Total N:		45		72	47	384	611
Contingency	Contingent	[.67]		[.70]	.87	.27	.13
	Assumed	[.49]		[.48]	.45	.52	.55
	Range				42	25	42
Stativity	Motion	.18		.83	[.62]	.28	[.38]
	Stative	.54		.44	[.48]	.52	[.52]
	Range	36		37		24	
Grammatical Person/Animacy	1 <sup>st</sup> Animate	.47		.68	[.44]	.39	[.56]
	2 <sup>nd</sup> Animate	.38		.12	[.55]	.81	[.45]
	3 <sup>rd</sup> Animate	.44		.47	[.46]	.55	[.54]
	3 <sup>rd</sup> Inanimate	.68		.50	[.69]	.43	[.31]
	Range	30		49		42	
Adv. Specification	No adverbial	[.48]		.44	.42	.56	.58
	Specific	[.35]		.77	.79	.31	.21
	Non-specific	[.67]		.61	.60	.34	.40
	Range			33	37	25	37
Type of Clause	Main	[.50]		.54	[.55]	[.48]	[.44]
	Subordinate	[.50]		.33	[.32]	[.59]	[.68]
	Range			21			

Indeed the data suggests that the latter explanation is the likeliest. As described in chapter 3, the twentieth century written data represents almost half century of language use, from 1961 to 1999. In contrast, the spoken data were collected over a much shorter period at the very end of the twentieth century, including the first years of this century: from 1991 to 2003. It seems reasonable to assume that the spoken data represents a more advanced

stage of the expansion of the *ir* + infinitive variant, a trend which is still emerging in the written data. The overall rate of usage (93%) and the almost total lack of conditioning demonstrate that the *ir* + infinitive periphrasis has very quickly expanded to fill all other contexts, except for those where the present tense is most firmly entrenched, contexts of adverbial specification and contingent eventuality. This rapid and near total expansion is consistent with a pattern of language change that has been observed in variationist studies:

At the earliest stage, the change apparently affects too small a population to serve as a model, but at some point it becomes perceptible, though usually beneath consciousness, and spreads through the community. (...) Once that point is attained, however, the change accelerates relatively rapidly toward the tailing-off point. (Chambers 2002: 361)

The validity of this pattern is confirmed by examining the direction of effect for relevant factors, even if the results were not selected as significant. Table 6.9 shows that, with the exception of grammatical person/animacy, all other groups of factors affect the present and the periphrastic *ir* + infinitive variants in the same direction with regard to speech and writing. Note that the constraint rankings of the factors are the same even when the groups of factors are not selected as statistically significant to the choice of the variants. For instance, *contingency* shows the same hierarchy of effect for speech and writing although it was not selected as statistically significant to the choice of the present variant in the twentieth century written data. *Stativity* and *type of clause* show the same hierarchy of

effect for written and spoken data, even though they are not statistically significant to the selection of the variants in speech anymore.

These empirical remnants of factors that once had a statistically significant effect confirm that the language represented by the playwrights is consistent with the patterning of vernacular speech. Thus, Table 6.9 presents empirical confirmation that these playwrights were using their vernacular speech. This result validates the written data of this study as representative of spoken Portuguese.

This finding was of fundamental importance for this study insofar as language change under investigation was expected to be most evident in vernacular, or language approximating the vernacular. In order to ensure the most appropriate written data for the analysis of language variation and change, it was critical that the source be representative of the vernacular speech. Therefore, a special amount of time and effort was dedicated to the characterization of the most trustworthy sources of the written data in Portuguese. This investigation was presented in detail in chapter 3. The search for the vernacular-like written data is now justified by the finding that they do indeed reflect natural language. The availability of these data offer insight into the most recent changes in the future temporal reference sector and help to illuminate the complex, but not random, process in which the variants are engaged.

The results show how the established variants gradually lose their favoured contexts, diminish in frequency and ultimately disappear. They also show how the new variants expanded their usage in the future sector. For example, each of the factors that are no longer selected as significant in Table 6.9 represents a potential context for the

expansion of the *ir* + infinitive variant. The next section will address the grammaticalization process of *ir* + infinitive in the data.

#### **6.6. The Trajectory of the Periphrastic Variant *Ir* + infinitive in the Variable Context**

This section focuses on the pathway *ir* + infinitive has followed in the variable context and whether this reflects its grammaticalization as future marker in the Portuguese language.

The original meaning of this construction corresponds to the lexical meaning of *ir* ‘to go’, i.e. physical movement to a deictic point different from speakers’, associated to a goal, e.g. *We’re going to Windsor to meet the King* (from Bybee, Perklins and Pagliuca 1994: 292). This movement toward a goal implied dislocation in space as well as time, which facilitated its association to a strictly temporal sense at a later stage (Bybee and Pagliuca 1987: 116; Bybee and Dahl 1989: 91; Heine, Claudi and Hünemeyer 1991: 173; Bybee, Perklins and Pagliuca 1994: 268, 292).

According to Bybee et al. (1994: 269), the expression of intention is also contained in the earlier construction insofar as: ‘when the speaker announces that s/he is going somewhere to do something, s/he is also announcing the intention to do that thing’; thus, the intentional meaning is not achieved through the grammaticalization process. The change takes place in later stages when the spatial meaning is lost in *go*-future, and prediction is added to the meaning of intention. Hence, the development of this construction would follow the following stages (source: Heine et al. 1991: 172):

Stage	Example	Prototypical Meaning
I	I am going to draw this...so that he can have a full picture.	Intention
II	It's going to be hot today.	Prediction
III	We're going to have a new mum.	Prediction

In the Portuguese corpus, the process from space/time to time is exemplified in (1) to (3):

- (1) [Aonde vais?]  
*Vou mudar* de camisa, que entendo, que estou mijado com alguma coisa mais. (Silva 1734. VE: Esopo: second part, scene 5: 122)  
 'Where do you go? I'm going to change my shirt, because I think, I'm wet with something else.'
- (2) Não posso responder, porque minha mãe já veio para casa, e lhe *vou falar*. (Silva 1734. VE: Filena: second part: scene 3: 108)  
 'I cannot answer it, because my mother came home and I'm going to talk to her.'
- (3) Olha, aqui em Sobradinho, eu acho que não. Se tiver reposição, *vai ser* esse sábado ou no próximo. (1992. Érica: A: 38)  
 'Look, here in Sobradinho, I don't think so. If they give us an extra class, it's going to be either this Saturday or next one.'

The examples in (1) to (3) show that there is, in fact, a transfer from space/time to time in the trajectory of the periphrastic form *ir* + infinitive. In the corpus, this space/time to time process is mainly established by the linguistic context, rather than attributing a single meaning to the construction itself. The transfer from intention to prediction is also established only by means of the occurrence of other linguistic elements in the structure. Such elements were investigated as conditioning factors in this study, which made it possible to assess the progress of this periphrasis along the pathway of grammaticalization

from verb of motion to future temporal expression. Some factors were examined as possible diagnostics to the degree of grammaticalization of the *ir* + infinitive variant: *temporal distance, grammatical person/animacy, type of clause and stativity.*

Tables 5.5 and 6.9 displayed the results of multivariate analyses of the influence of these factors on the selection of the *ir* + infinitive variant in the nineteenth and twentieth centuries. Recall that *ir* + infinitive was not considered as form of future temporal reference when dislocation in space was conveyed by the context, as in (1) above. Thus, all tokens of *ir* + infinitive examined in this study showed a temporal reading.

In Tables 5.5 and 6.9, there do not appear to be any restriction on the distribution of this future variant, as one might expect from a strictly lexical, spatial reading of the construction. Rather it is only disfavoured in certain contexts. A close examination of the contexts offers evidence in favour of its grammaticalization, as shown below:

(i) The development of *ir* + infinitive into future marker is correlated to its ability to express *distal future*, since the usage of a form specifically for immediate future is considered to precede a pure future meaning in an evolutionary representation. Thus, this periphrasis is first expected to denote immediate future, and only later to extend its usage through events and actions that are not totally anchored in the immediate future.

Results from the nineteenth century (Table 5.5) show that *temporal distance* is statistically significant to the selection of *ir* + infinitive, with proximal contexts favouring this variant. In the twentieth century, however, this effect is no longer evident, exactly as

would be expected if the variant was evolving from an expression of near futurity to a more generalized future marker.

(ii) This periphrasis is also said to be first expressed with subjects capable of volition and to reflect speaker's attitude (cf. Poplack and Tagliamonte 2000: 334-5). Thus, human and first person subjects are the primary locus for its usage. Only in a later stage of its development, 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 and Pagliuca 1987: 109-10), would make possible the association of this future form with inanimate subjects. The generalization of *ir* + infinitive to non-first person and inanimate subjects would be indicative of grammaticalization.

Results from the nineteenth century (Table 5.5) show that non-first person and inanimate subjects are associated with this periphrasis in the nineteenth century. In that century, grammatical person/animacy is not selected as statistically significant to the choice of *ir* + infinitive. This result suggests that the grammaticalization of *ir* + infinitive was sufficiently advanced at that time. The effect of this group of factors is, however, unclear in the written data of the twentieth century, and only later in contemporary spoken language, is it clear that it does not determine the usage of this periphrasis in Portuguese (Table 6.9). Thus, in contemporary Brazilian Portuguese, *ir* + infinitive is advanced in the process of grammaticalization with regard to the *grammatical person/animacy* factor.

(iii) Generally, the use of future forms in *subordinate clauses* indicates attainment of an advanced stage in the evolutionary process; in other words, only mature future forms reach such clauses. As stated by Bybee et al. (1994: 244) '(...) because late-developing uses, such as subordinate clause uses, are characteristic of the most mature future forms and are not displayed by younger future forms (...)'. However, *ir + infinitive* is also said to be first originated in subordinate clauses (cf. Danchev and Kytö 1994: 68, as cited by Poplack and Tagliamonte 2001: 226), in such a case its expansion to main clauses is expected in the process of grammaticalization. In order to find out the original syntactic location of *ir + infinitive* in Portuguese, and to examine whether this variant expanded its usage beyond this location, the group of factors *type of clause* was tested in this study.

Since the nineteenth century, when *ir + infinitive* was first examined quantitatively, it has been used in both clause types, but has shown a higher probability of usage in subordinate contexts (cf. Tables 6.5 and 6.9). However, *type of clause* has never been selected as statistically significant to its occurrence. The effect of subordinate clauses on the realization of this periphrasis was also found in English (Poplack and Tagliamonte 2001: 228). In both languages, the periphrasis has shown a higher probability of usage in subordinate clauses, but this syntactic context is not significant to its choice. The results suggest that *type of clause* is not as relevant to determine the stage of grammaticalization of this periphrasis as, for instance, proximity and verbs of motion.

(iv) Finally, the degree of grammaticalization should affect the propensity of *ir* to occur with other verbs of motion. The reasoning that motivates this statement is that the occurrence of the verb of motion *ir* with another verb of motion indicates desemanticization of its lexical content (cf. Poplack and Tagliamonte 2001: 230).

In contemporary spoken language, the group of factor *stativity* is not statistically significant to the selection of this periphrasis (cf. Table 6.9). As expected in the process of grammaticalization, *ir* + infinitive has advanced to the point where it is used more often with verbs of motion in spoken Brazilian Portuguese than it was in the written language. This result confirms the advanced stage of grammaticalization of this periphrasis in the language.

Recall, however, that the class *verbs of motion* was statistically significant to the choice of the present variant from the sixteenth and eighteenth centuries to the written twentieth century data. Furthermore, the main verb 'to go' has been almost categorically associated to the present variant in the language. From this follows that in Brazilian Portuguese the association of *ir* + infinitive with the main verb 'to go' is typically avoided. This restriction can be attributed to the widespread prescriptive belief that *vou ir* 'I'm going to go' is redundant. This retention of the original lexical sense of 'to go' suggests that the process of grammaticalization has not yet gone to completion in spoken Brazilian Portuguese. In order to achieve the farthest stage in the grammaticalization path, *ir* + infinitive still needs not to be disfavoured with the main verb 'to go'.

On the basis of these findings, it seems reasonable to assume that the periphrastic *ir* + infinitive variant has been undergoing increasing grammaticalization as a form of future temporal reference in Brazilian Portuguese, at least since the nineteenth century. It has expanded its original usage in this process; as a result, in contemporary language factors of *temporal distance*, *grammatical person/animacy* and *stativity* no longer play a significant role in determining its selection in the future context.

As the multivariate analyses confirmed, the grammaticalization process is not only related to the increase in frequency. The trajectory of *ir* + infinitive in the cline of grammaticalization can be related to the selection of factors that constrain its choice in the future temporal context. Thus, we can confirm empirically the generalization of this periphrasis in the future sector. According to Bybee et al (1994):

(...) in generalization the context is important not because meaning from the context is transferred to the developing gram, but because use in more contexts causes a shift in the primary message being conveyed, which contributes eventually to the loss of specific features of meaning. (Bybee et al. 1994: 292)

*Ir* + infinitive expanded its usage from movement toward a goal, as in example (1), to future marker. It has weakened its specific meaning, and as a future marker occupies linguistic contexts that once favoured the selection of other future variants. Throughout the centuries, this variant has expanded in its use to the point that in contemporary speech it maintains vestiges of its movement meaning only when used with the main verb 'to go'.

The final step in the grammaticalization process of this periphrasis is to be productively used in such a context, demonstrating the complete loss of its lexical meaning.

### **6.7. Prescription vs. Usage**

Having uncovered the conditioning factors responsible for the partitioning of the future temporal reference sector in contemporary and historical Portuguese, it merely remains to ask: is there any relation between prescription and natural language usage in the future temporal reference sector in Portuguese? Based on the findings of this research the answer appears to be “no”.

The summary of grammatical prescriptions presented in chapter 2 (cf. Table 2.2), and the results of the multivariate analyses reported in chapters 5 and 6 of this study confirm that, with a few exceptions, there is no regular correspondence between the prescribed use of the variants by grammarians and their actual distribution in writing and speech. Not only do grammarians appear to have failed to recognize the real role of grammatical factors in influencing the selection of the variant, but they also appear to have remained oblivious to dramatic changes in the overall rates of the variants, still attributing to the synthetic future a prominence that is now several centuries out of date. Unable to quantify the complex changes that have taken place, grammarians are often forced to fall back on vague distinctions that attribute different meanings or intentions to different variants.

Some concluding words are given on prescription and usage in this section. For that, it is focused on: (i) the meanings attached to the future forms by grammarians; (ii) the

distribution of the future variants in formal/informal and written/spoken data; and (iii) the results of speakers' level of education.

#### 6.7.1. Meanings attached to the future variants

Although the synthetic future is prescribed as the correct future form, variation in the future sector has been attested in the prescriptive literature since the publication of the second grammar book of Portuguese in 1540. Variation, however, is anathema to prescriptive grammarians and, forced to acknowledge what they could not explain, they were pressed to provide principled explanations for the existence of multiple forms. For the most part, the selection of the forms has been related to speakers' intent, entailing the interpretation of specific meanings for each form. The synthetic future is the sole exception, since it has always been prescribed as the default future temporal referent, and only in the twentieth century grammars it started to be regularly associated with other meanings than future time (cf. Table 2.2).

The attempt to attach a unique function to each form fails on a number of levels. Failure to properly circumscribe the variable context results in confusion between future uses of the variant forms and related forms that might have other possible readings, such as *command*, *obligation*, or *movement*. Furthermore, there is some evidence that grammarians failed to recognize emergent uses of new future forms, creating only a partial picture of the future paradigm (cf. Table 2.2). Even assuming that consensus can be reached on the relevant forms and the meanings ascribed to them, the correspondence between the two is somewhat haphazard: the same meanings are associated to the present and the periphrastic variants *haver* + infinitive and *ir* + infinitive, e.g. *necessary occurrence*, *proximal*, *certainty*, *firm intention*, *resolution*, *determination*, *emphasis etc* (cf. Table 2.2).

Furthermore, if such considerations of meaning were relevant to variant selection, then it is expected that they would mitigate against the clear effect of straightforward grammatical factors such as, *polarity, verbs of motion, contingency and adverbial specification*, all of which were selected as statistically significant to the occurrence of one or other variant, but are mostly ignored in the literature (cf. Table 2.2).

The most significant findings of this study are (1) the finding that the variation in the selection of future forms in Portuguese is regular and rule governed, and (2) that the future reference paradigm has undergone a radical change whereby the synthetic future has retreated from its traditional contexts and almost disappeared in twentieth century language. Concomitant with this decline of the synthetic future, the periphrastic *ir* + infinitive variant reaches the overall probability of occurrence of .766 in the written data, and .929 in speech in the twentieth century. Regardless of this, the prescriptive literature of the twentieth century persists in affirming that the synthetic future is the expression that conveys future time in the Portuguese language, and it still associates the usage of the periphrasis to special meanings, as for instance, *proximity* and *certainty*.

The unsystematic and very often long-term replicated comments of the Portuguese prescriptive literature do not reflect the usage of the future forms in the language. Clearly the speed of language change has outstripped the pace of prescriptive adaptation to the reality of language, if indeed the prescriptive tradition ever succeeded in describing the facts of real language.

### 6.7.2. Comparison of the Distribution of the Variants in Formal/Informal and Writing/Speech

This subsection compares the distribution of the variants in the twentieth century data with that of Santo's and Baleeiro's studies (cf. chapter 4). Since the distribution of the variants in their data was quite different from that of the present study, it was determined to verify whether this dissimilarity was due to a difference in the sources (cf. Bortoni-Ricardo 1998). As it turns out, the difference in the distribution is best explained as a difference in medium. Specifically it appears that only the formal written language retains the synthetic future as the norm in Brazilian Portuguese.

Table 6.10 shows that the distribution of the variants varies according to formality and medium, particularly with regards to the selection of the synthetic future. The synthetic future is maintained as the preferred choice in the formal writing (73%), while it is virtually absent from the informal speech (0.5%).

**Table 6.10: Distribution of future variants according to three studies with different corpora in twentieth century Brazilian Portuguese.**

VARIANTS	Formal Writing Santos (1997)	Formal Speech Baleeiro (1988)	Informal Writing (From this study)	Informal Speech (From this study)
<i>Ir</i> + infinitive	22%	77%	77%	92.5%
Present	2%	9%	14%	7.0%
Synthetic future	73%	13%	9%	0.5%
<i>Irei</i> + infinitive	3%	1%	---	---

The synthetic future and the periphrastic *ir* + infinitive show a reversal in usage in formal writing and in informal language/speech, what suggests that the synthetic future is

associated to the formal writing variety. This distribution also shows that both formal speech and informal writing do not differ in terms of preference for *ir* + infinitive usage (77%). Accordingly, we notice that this periphrasis has a higher percentage of usage in speech and informal language, being virtually categorical in the informal spoken data (92.5%). This appears to be directly related to the unavailability of synthetic future in this context.

This comparison shows that the phenomenon under investigation in this study follows a parameter informal-formal and a parameter speech-writing in the language. Both parameters may be theoretically independent, but they are interconnected in usage. However, the most unusual pattern is that of the formal writing, which is represented by a variant that has practically abandoned the future sector.

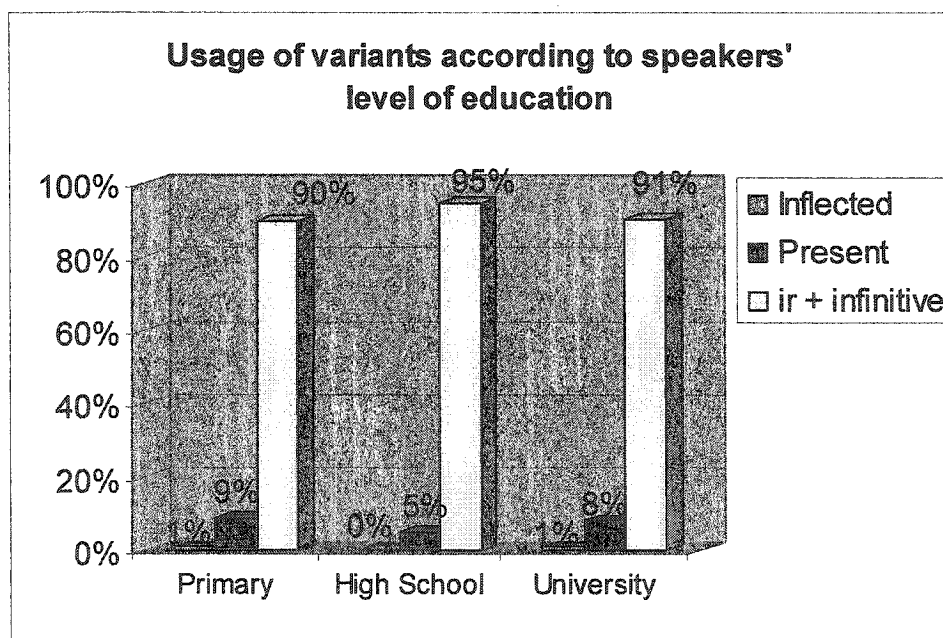
In order to understand the importance of this comparison to the effect of prescriptive rules in Brazilian Portuguese, recall that the rates of the synthetic future in plays of the sixteenth and eighteenth century were 66% and 57%, respectively (cf. Table 5.2). Thus, in Brazilian Portuguese, the rate of the synthetic future in the twentieth century formal writing is actually higher than it was in the sixteenth and eighteenth century Portuguese comedies, despite a general and rapid decline in the use of the synthetic variant overall. This fact stresses how Brazilian formal writing adheres to a prescriptivism that lags far behind actual language use. It represents prescriptive beliefs about the way language should be used, becoming artificial and totally detached from natural language.

### *6.7.3. Level of Education*

One question raised during the first steps of this study was to what extent education would affect the usage of the future variants in spoken Brazilian Portuguese. As discussed

in Chapter 2, the synthetic future is the prescribed or standard form for future temporal expression in the Portuguese literature. Therefore it was expected that speakers with a higher level of education, and presumably more exposure to prescriptive influence, would show a higher rate of the standard variant in their speech. As the results in section 6.4 already showed, this hypothesis was not supported by the data. The synthetic future is almost absent from the spoken data, and *ir* + infinitive is almost categorically selected in future contexts, regardless of level of education.

The graph below shows the percentage of usage of each variant in the informants' speech according to their grade in Brazilian educational system. Notice that the percentage of *ir* + infinitive is almost the same at all three levels: primary (90%), high school (95%) and university (91%).



*Figure 6.1: Speakers' level of education*

Note also that the primary-level students show a slight preference for the present variant when comparing its percentage of usage to that of the other two groups. In addition, the high-school level speakers show the highest usage of *ir* + infinitive; even though, the difference is in fact minor (see chapter 3 for details on sample formation). Relevant to this study is the scarce number of the synthetic variant found in the speech of university-level speakers.

What does this distribution show us about the expression of future temporal reference in Brazilian Portuguese? Has the prescriptive pressure in favour of the synthetic form as the standard future failed? Has the prescriptive tradition not successfully echoed through the society? Do Brazilian public schools ignore prescriptive orientation or do they effectively implant standard language? Why has *ir* + infinitive prevailed throughout the centuries in Portuguese, despite prescriptive injunctions to the contrary?

To properly answer all these questions and some more that may naturally be derived from them, it would be appropriate to carry out empirical studies on language teaching in Brazilian schools. Nevertheless, the findings of this study make it possible to reflect on some issues that have been raised in the literature with regards to prescriptive tradition and standard language (see Bortoni-Ricardo (in press); Cheshire 1999; Chambers 1995; Hickey 2000; Lippi-Green 1997; Milroy 1999; Milroy 2000; Milroy & Milroy 1999; Poplack 2003; Trudgill 1999; Watts 1999; Watts 2000).

First, it is worth noting that *ir* + infinitive and the present variants are not stigmatized in Portuguese. Thus, although the synthetic future was elected the standard future form by grammarians, there was no stigma attached to the use of the other forms. Hence, the alternate forms have the potential to be legitimized in the grammars (cf. chapter

2), unlike other linguistic phenomena that have run afoul of prescriptive rules for example, nominal and verbal agreement.

According to Poplack (2003), this pattern is also found in English and French prescriptive tradition. She posited that grammarians' efforts to neutralize inherent variation in language are characterized by one or more of the following attitudes: (i) totally ignoring the variation; (ii) condemning the non-standard variants; or (iii) avoiding the form-function asymmetry by establishing different functions or readings to each form. In the future temporal sector in Portuguese, grammarians follow these stages: (i) first they disregard any future form usage as a temporal usage, but the synthetic future; (ii) then they acknowledge the other variants; (iii) however, in that case they impute to them numerous readings; (iv) finally, a few start to address the usage of alternate forms as frequent in *spoken language, informal speech, conversation* etc. (cf. chapter 2 and Table 2.2).

In fact, the synthetic future has always been assumed as the standard future form in the standard Portuguese language. The Portuguese prescriptive tradition keeps working in favour of this idealized language. As remarked by Trudgill (1999: 125)<sup>38</sup> and Milroy (2000: 13), the standard language is a social construct that is mainly characterized by its grammatical forms and by invariance, which is enhanced by Poplack's (2003) findings that grammarians have worked to neutralize variation in language throughout the centuries.

Despite prescriptive efforts (cf. chapter 2), the findings of this study show that grammarians and the school system have not affected the course of natural language, i.e. the vernacular. The school system, which perpetuates prescriptive rules and promotes the

---

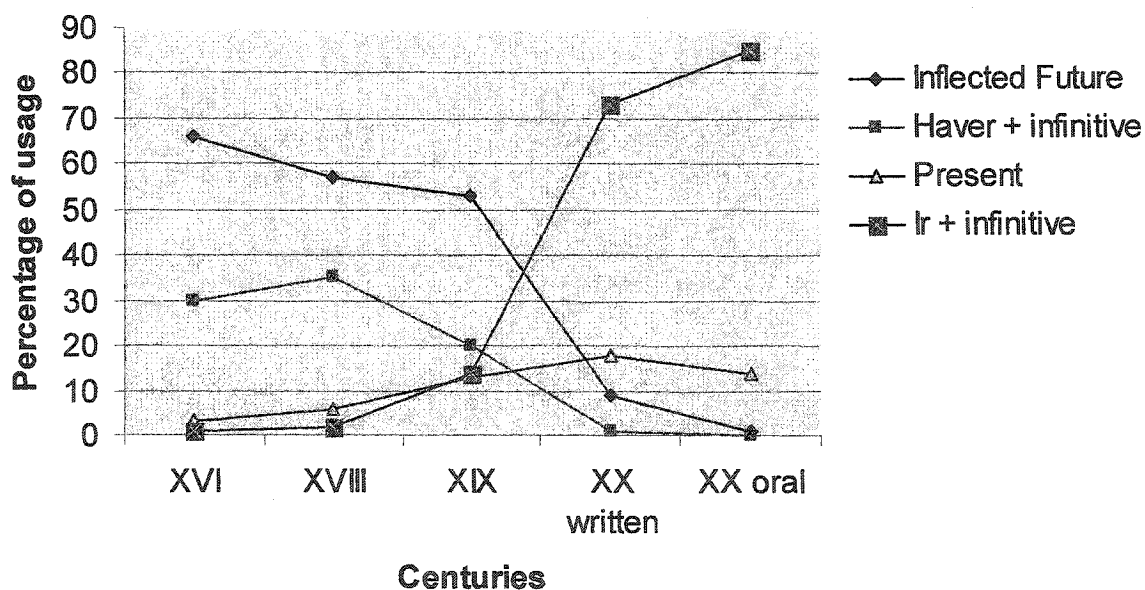
<sup>38</sup> I am aware that for Trudgill (1999:125) "these grammatical forms are not necessarily identical with those which prescriptive grammarians have concerned themselves with over the last few centuries"; however, in Brazilian Portuguese the only difference between both standard grammatical forms and prescriptive grammatical forms is related to non-stigmatized forms, which are used as in a *continuum*; otherwise, they are exactly the same forms because stigmatized variants are not allowed in the standard language.

standard language, seems to have been effective on the maintenance of the standard future form only in the formal written language. Brazilians who are able to achieve a high educational level do not transfer an invariant synthetic future usage to their vernacular. On the contrary, the results of this study show that they do not use the synthetic future. In spoken language, the norm is the periphrasis. The findings of the quantitative analyses revealed that *ir* + infinitive has succeeded in the future sector. Most important to this discussion, they demonstrated that the variants have been involved in a complex process of change throughout the centuries in the language, and that this process has been mainly ignored by the prescriptive tradition.

## Chapter 7 Discussion and Conclusion

---

The present study has examined future temporal reference in the Portuguese language. Making use of standard variationist and comparative methodology, the preceding chapters have shown that the different forms of future temporal reference have been in variation for centuries in the language, and, most interestingly, they have been involved in a complex process of incursion and withdrawal from the variable context, which ultimately changed the representation of the future sector in contemporary Brazilian Portuguese, as illustrated by the graph below:



*Figure 7.1: Usage of the future variants throughout the centuries in Portuguese*

This study has shown that the difference in the frequency of each variant in the different centuries is correlated to their displacement of or expansion into some specific linguistic contexts. The variationist methodology has made it possible to pinpoint these linguistic contexts and to identify how the 'new' future variants, i.e. the present and *ir* + infinitive, gradually occupied contexts that earlier favoured the occurrence of 'old' variants, i.e. the synthetic future and *haver* + infinitive. The identification of these linguistic contexts, here tested as discrete factors, uncovered the direction of change in the Portuguese future sector. As stated by Labov (2001):

The changes that have most profound effect and represent the main stream of linguistic evolution are not isolated shifts of single elements, *but movements of one or more elements in a continuous direction.* (Labov 2001: 421) [emphasis added]

In order to establish the direction of the linguistic evolution in the future sector as well as the factors that constrained this evolution, this study followed specific stages. 1. Remarks on usage and/or meanings of the future forms mentioned by historical, descriptive and prescriptive grammarians were traced as far back as the sixteenth century. 2. The object of this investigation, the future temporal reference context, was circumscribed. 3. Diachronic (written) and synchronic (written and spoken) data were gathered and transcribed. 4. The prescriptive remarks were operationalized as factors in order to systematically account for the phenomenon and to test the linguistic hypotheses of this research. 5. The variable was examined by century and the results of the quantitative analyses were presented by variants, which made it possible not only to assess the factors

that were favouring each variant, but also to identify the factor that influenced the most each variant in each century and throughout the centuries in the language.

Besides the confirmation of change, the results of this study showed that prescriptive remarks do not explain the usage of the future temporal referents in the language. In addition, the results of this study provide strong support for the usage of written data (popular plays) in reconstructing early stages of a language in order to diachronically investigate variation. Particularly, this study endorses the value of written data to the analysis of non-stigmatized variants.

This chapter reviews the findings of Chapters 5 and 6, summarizes the findings for each variant of the future temporal reference, and considers the implications of these findings to the study of the future sector and language change.

## **7.1. Findings of this Study**

### *7.1.1. The Future Variants*

The findings of this study regarding the usage of the future forms in the variable context throughout the centuries were presented and discussed in chapters 5 and 6. They provided support for the conclusion that the representation of the future temporal reference has undergone change in the language. Moreover, the selection of the statistically significant conditioning factors in the quantitative analyses made it possible to trace this change throughout the centuries. More interesting for the understanding of the variation and change in the future sector is the fact that in each century it was possible to identify the precise linguistic contexts in which the variants lost or gained ground as well as the way

this process became decisive for their maintenance in the future sector. The major findings are presented in the next subsections.

#### 7.1.1.1. *The Synthetic Future*

The behaviour of the synthetic future in future temporal reference contexts has been the focus of great interest in this study. Since the sixteenth century (cf. Table 2.2), this variant has been acknowledged by prescriptive grammars as the form that expresses futurity in Portuguese. Its status as the default variant is also noticed in French (cf. Poplack and Turpin 1999). However, in both languages prescription does not conform to the actual usage. Particularly in Portuguese, this research has shown that since the sixteenth century the synthetic future has occurred as only one of the options of futurity, and since then has retreated in the language.

In the sixteenth and eighteenth centuries, this variant had an overall probability of occurrence of .622 in the language. Future temporal reference was also expressed by *haver* + infinitive (.329) and to a lesser extent by the present variant (.049). The strongest linguistic constraint on the choice of the synthetic future at that time was *contingent events/actions*, followed by *stative verbs*, *statements* and *affirmative utterances*. Interestingly, the selection of contingent context as conditioning the usage of the synthetic future in the sixteenth and eighteenth century Portuguese is also found in contemporary spoken French (Poplack and Turpin 1999: 149).

However, this study showed that in Portuguese, unlike French, the synthetic future lost this favouring context in the nineteenth century. The multivariate analyses made it possible to detect the first sign of an ongoing change in the future sector, i.e. the

displacement of the synthetic future from the linguistic context that earlier contributed the most to its usage.

In the nineteenth century, the synthetic future had an overall probability of occurrence of .545 in the language. Three other future variants productively occurred in the future sector: *haver* + infinitive (.211), the present variant (.098) and *ir* + infinitive (.146), the latter being incipient. As already mentioned, *contingent events/actions* were no longer statistically significant to the occurrence of the synthetic future; this context began to favour the present. Another significant loss of ground was detected in the *polarity* group of factors. The synthetic future was no longer favoured by *affirmative* utterances. This context became the strongest linguistic constraint to the occurrence of the incipient *ir* + infinitive. In contrast to those contexts where the synthetic future declined, *stative* verbs continued to favour the synthetic future in the nineteenth century.

The most significant contributing factor in the nineteenth century was the presence of *non-specific adverbs*. An adverbial specification pattern emerged in Portuguese in the nineteenth century: *non-specific adverbs* favoured the synthetic future; *specific adverbs* favoured the present variant; and *no-adverbial* favoured *ir* + infinitive. These effects were consistent with results found to contemporary spoken French (Poplack and Turpin 1999: 149). Given the time period of emergence of this pattern in Portuguese, it would be interesting to investigate whether in French the adverbial specification condition also surfaced in the nineteenth century. Such confirmation would suggest that the adverbial specification pattern is not language specific, and would reinforce the need of such an investigation in other Romance languages.

In the twentieth century the usage of the future variants was investigated in written data from the middle of the century onwards and in the spoken language of the end of the century. In the written data, the synthetic future had the overall probability of occurrence of .090. Other two variants occurred productively in the corpus: the present variant and *ir* + infinitive. The synthetic future continued to be favoured by stative verbs, which at that time were the most significant contributing factor to its selection. Moreover, it was also favoured by inanimate subjects. Both factors are best interpreted as retaining the synthetic variant.

The multivariate analyses made it possible to identify another linguistic context lost by the synthetic future in the written twentieth century. This variant was displaced from the non-specific adverbial context by the present variant. As seen in this study, the synthetic future not only reduced its frequency throughout the centuries in the language, but it gradually lost its most significant contributing factors as the present and *ir* + infinitive variants expanded their usage in the language. In addition, the results showed that stative verbs started to be selected as statistically significant to the choice of *ir* + infinitive. Such result leads us to foresee that the synthetic future will also decline in this context, which has favoured its occurrence since the sixteenth century.

As attested by this evolutionary process, the synthetic future moved away from the future sector in Portuguese. In the spoken data of the end of the century, this variant is virtually non-existent. It accounted for only one per cent of the data. This study also showed that speakers ranging from primary to university level of formal education did not show any difference regarding the usage of the synthetic future in the spoken language. Furthermore a comparison of the distribution of the written and spoken twentieth century

data of this research with those of two other studies in formal Brazilian Portuguese confirmed that the synthetic future is only maintained as a future marker in highly formal written language.

#### 7.1.1.2. *Haver + infinitive*

Since the sixteenth century, this variant has been acknowledged by prescriptive grammars as expressing mainly other meanings than futurity in Portuguese (cf. Table 2.2 and Figure 2.1). However, this study confirmed that this periphrastic construction was used in the future temporal context in the language. Moreover, its usage was conditioned by specific factors throughout the centuries.

In the sixteenth and eighteenth centuries, *haver + infinitive* had the overall probability of occurrence of .329. The future temporal sector at that time was also represented by the synthetic future and the present variant, the latter being the incipient variant in the plays. *Haver + infinitive* was strongly favoured by *assumed events/actions*. This factor continued to have the greatest influence on the selection of this variant in the nineteenth century.

During the nineteenth century, the future temporal reference context was represented by other three variants: the synthetic future, the present variant and *ir + infinitive*. The overall probability of *Haver + infinitive* decreased (.211). It was observed that some factors, such as affirmative utterances, subordinate clauses, stative verbs and no-adverbial specification, revealed no statistically significant effects on this variant. The direction of effect showed, however, that these factors were favourable to this variant, as well as to *ir + infinitive*. It was concluded that the preferred occurrence of *ir + infinitive* in

the contexts that were also preferred by *haver* + infinitive was evidence that *haver* + infinitive was being superseded in the future sector. Multivariate analysis shows that *haver* + infinitive was excluded from the adverbial specification pattern that emerged in the nineteenth century. This result was interpreted as a further sign that *haver* + infinitive would not last in the future sector. In the written twentieth century language, this variant lost the *assumed context*, which had favoured its occurrence since the sixteenth century, further loosening its hold on the future sector.

#### 7.1.1.3. *The Present Variant*

The present variant revealed a steady occurrence in the future sector throughout the centuries, as shown in this chapter by Figure 7.1. The usage of this variant showed a slight increase as the usage of the synthetic future decreased in the language.

The first and valuable finding of the analysis of the present variant was the almost categorical preference of the main verb ‘to go’ for this variant in the future context. Due to this lexical association, this verb was excluded from the quantitative analyses from the nineteenth century onwards, when this correlation was effectively verified. In fact, in contemporary spoken Brazilian Portuguese, ‘to go’ upholds the usage of the present in the future sector.

In the sixteenth and eighteenth centuries, this variant had an overall probability of occurrence of .049 in the language. Hence, it was incipient in the corpus. The strongest linguistic constraint on its choice at that time was *verbs of motion*. This effect was observed throughout the centuries, except for spoken language. By the range of the group of factors, the multivariate analyses enabled us to follow the lessening of this factor: (i) range of 49 in

the sixteenth and eighteenth centuries, (ii) range of 45 in the nineteenth century, (iii) range of 37 in the written twentieth century, and (iv) range of 14 in the spoken twentieth century, when it was no longer selected as statistically significant to the choice of the present variant.

The evolution of this variant in the future sector showed that it gradually became favoured by the most significant contributing factors of the synthetic future. In the nineteenth century, the present variant was favoured by *specific adverbs* and *contingent contexts*, besides *verbs of motion*. As already mentioned in section 7.1.1.1, contingent contexts were earlier selected as statistically significant to the occurrence of the synthetic future.

Another noteworthy finding was the effect of specific adverbs on the selection of the present, which was revealed as soon as the adverbial specification pattern emerged in the language in the nineteenth century. The effect of specific adverbs was interpreted as a disambiguating effect in the language, since the usage of the present variant could be understood as referring to present temporal reference instead of future one if no extra information was given by the context.

In the written twentieth century, this variant reached the overall probability of occurrence of .144. This increase did not correspond only to a greater frequency in the corpus. As shown by the multivariate analyses, *non-specific adverbs* started to be selected as statistically significant to the choice of the present. Recall that non-specific adverbs were the most significant contributing factor to the selection of the synthetic future in the previous century. Thus, this study showed that the present expanded its usage through a new linguistic context.

It is worth noting that the contribution of specific and non-specific adverbs to the selection of the present variant finds parallel in spoken French (Poplack and Turpin 1999: 149). The difference between Brazilian Portuguese and Canadian French seems to lie on a more advanced stage in the process of change of the former, since adverbial specification was no longer significant to the selection of the synthetic future in written twentieth century Brazilian Portuguese, and the synthetic future is no longer productively used in the spoken language (only 1%).

The last valuable finding of this study with respect to the occurrence of the present variant was provided by the results of the analysis of the spoken twentieth century data. In contemporary spoken language, future temporal reference is expressed by the present variant (.071) and *ir* + infinitive (.929). As shown by this study, the present is actually employed in a very restricted context. It is retained in spoken Brazilian Portuguese by *contingent events/actions* and *adverbial specification* (specific and non-specific adverbs).

#### 7.1.1.4. *Ir* + infinitive

The situation described in this study provides substantial evidence that the periphrastic *ir* + infinitive variant, far from being reserved for the expression of *movement*, *proximity*, *intention*, *resolution*, *certainty* or *necessity* (cf. Table 2.2), is the form that mainly expresses futurity in twentieth century Brazilian Portuguese. This study has shown that the process of evolution of this periphrasis is marked by its constant expansion in the future sector since the sixteenth century when it accounted for only one per cent of the data.

This periphrasis was quantitatively first analysed in the nineteenth century when its overall probability of occurrence was .146. In that century, other three future exponents

productively occurred in the future sector: the synthetic future (.545), *haver* + infinitive (.211) and the present variant (.098). The multivariate analysis showed that the occurrence of this periphrasis was constrained by three factors: *affirmative utterances*, *no adverbial specification* and *proximal future eventuality*, being affirmative utterances the most significant contributing factor.

The effect of those factors on the choice of *ir* + infinitive reveals a parallel to the results found for spoken Canadian French (Poplack and Turpin 1999: 149). In both languages the periphrasis is favoured by those same factors, following the same hierarchy of constraints (as shown by their ranges). The effect of polarity (affirmative/negative utterances) is remarkable: *ir* + infinitive is highly avoided with negation in both languages. A notable difference, though, is that in Brazilian Portuguese those results are for the written nineteenth century data, while in Canadian French these results are for contemporary spoken language. Such a difference in time period and similarity in the hierarchy of constraints provide further evidence that change in the future sector is advanced in Brazilian Portuguese.

In the written twentieth century, this periphrasis had an overall probability of occurrence of .766 in the language. Two other future variants productively occurred in the future sector: the present variant (.144) and the synthetic future (.090), the latter receding. The multivariate analysis made it possible to identify two linguistic contexts in which *ir* + infinitive gained ground: (1) first, this variant was favoured by *assumed events/actions*, which earlier maintained *haver* + infinitive in the future sector; (2) second, it started to be favoured by *stative verbs*, which was the most contributing factor to the selection of the synthetic future in the written twentieth century data.

This evolutionary process made it clear that *ir* + infinitive did not only increase in frequency in the data, but gradually displaced other variants from the future sector. In spoken Brazilian Portuguese, it reached the overall probability of occurrence of .929, which revealed that the ongoing change is near-completion. The only other variant productively used in the spoken language is the present (.071).

At this stage, factors that are considered to be related to the primary meaning of this periphrasis are no longer revealed as statistically significant to its occurrence. In chapter 6, these factors were examined, as well as the pathway of grammaticalization that *ir* + infinitive followed in the data. In general, the findings of this study seem to be compatible with the pathways of grammaticalization established in the literature.

The multivariate analysis thus afford us to ascertain that contingent actions/events and contexts with adverbial specification are the only two linguistic contexts that still favour the occurrence of the present variant in the language. As mentioned in chapter 6, those two contexts and the future of the main verb 'to go' are the last barriers to *ir* + infinitive. They remain as the last barriers to the completion of linguistic change.

### 7.1.2. *The Value of Plays*

The results of this study call attention to the value of popular plays, in particular comedies, in reconstructing earlier stages of language. In the case of the representation of future temporal reference, these written sources made it possible to examine the variation, and the process of change this representation has undergone throughout the centuries.

Chapter 3 presented the search for written data considered to be representative of speech. In this search, sociohistorical research supported the idea that some playwrights

were regarded as popular; that is they represented the speech and customs of ordinary people. However, the most important criterion observed in the selection of the corpus was the ability of those playwrights to reproduce linguistic features others than the future forms that could characterize their work as colloquial, such as double negation, non-standard pronoun placement, phonological reduction and second person verbal non-agreement.

Quantitative analyses of the diachronic data showed that the playwrights were able to represent the evolution of the variants in the future sector not only in terms of their overall rates but in terms of their conditioning. Most important to this research was the finding that the plays showed a consistent pattern of conditioning and hierarchy of constraints, such as the gradual process of displacement of the 'old' variants in certain contexts as the present or *ir* + infinitive variants gradually expanded into them (cf. section 7.1.1). This coherent evolutionary process supports the conclusion that the playwrights used the future variants in the plays as in natural language.

The multivariate analyses also revealed parallels between the results of the written and the spoken data of the twentieth century. These parallels cover conditioning that was shared (contingent contexts and adverbial specification) by the present variant. Moreover, according to the literature on grammaticalization, the contributing factors selected as significant to the written data and not selected for the spoken, represent proper linguistic contexts to the expansion of *ir* + infinitive. This expansion suggested that the spoken data from the end of the century portrayed a further stage in the process of change (cf. Chapter 6). This study has shown that the consistency of the evolutionary process of the variants is empirical evidence of the validity and reliability of the plays for variationist research.

### 7.1.3. Prescriptivism

Throughout this study, the analyses have provided evidence that much of what is prescribed regarding the use of the future forms does not correspond to their actual use in the language. The findings of the present study with respect to the use of the *synthetic future*, the *present variant* as well as the periphrastic variants *haver + infinitive* and *ir + infinitive* indicate that all four future forms have occurred in future contexts in the Portuguese language for centuries. Far from being the future form *per se*, the synthetic future dominated future reference with an overall probability of occurrence of .622 only in the sixteenth and eighteenth centuries. In the nineteenth century, all four variants were productively used in the future sector, and in the twentieth century, the periphrastic variant *ir + infinitive*, which is mainly associated to the expression of *proximal future* and *certainty* in prescriptive literature, was used instead as the norm, reaching the overall probability of occurrence of .929 in Brazilian Portuguese vernacular speech.

The above line of evolution points to the fact that the actual use of the forms of the future temporal reference and prescriptive instructions has been unrelated for centuries in the Portuguese language (cf. Table 2.2). The present study addressed this incongruity in Chapters 2, 5 and 6.

With the exception of a few remarks, such as the observation of the co-occurrence of adverbs of time with the present variant (e.g. Cunha & Cintra 1984/ 1999: 438, Thomas 1969: 116 and Van Achter et al. 1996: 155), the attachment of a *positive* meaning to the synthetic future in 1852 by Soares Barboza, and comments on the infrequent use of the synthetic future in contemporary spoken language (e.g. Cunha & Cintra 1984/1999: 459; Kury 1989: 136; Thomas 1968: 123; 1974: 117; Van Achter et al. 1996: 155), prescriptive rules have not succeeded in indicating the favoured linguistic contexts of each variant in the language.

As mentioned in Chapter 2 and later confirmed by the quantitative results, this incongruity is a consequence of grammarians' efforts in establishing a form-function symmetry in the future sector. This form-function symmetry cannot account for the dynamic process of variation and change examined in this dissertation.

The divergence between prescription and use is nowhere better demonstrated than in contemporary Brazilian Portuguese, in which grammarians, with a few exceptions, still prescribe the usage of the synthetic future and associate meanings to the other future variants without considering that *ir* + infinitive is the preferred choice in formal speech, informal writing and vernacular language, being almost the only selected variant in the latter.

## **7.2. Conclusion**

What are the implications of these major findings to the analyses of variation and change? First, they provided support to the application of the variationist theory and the comparative method in diachronic analyses. Second, they revealed that, even in a complex

process of change, in which the number, type and frequency of variants are different in all centuries, it is possible to follow the conditioning factors of the variants, to state their direction of effects, confirm the maintenance of the hierarchy of constraints, and foresee the withdrawal of the variants from the linguistic context. Particularly for the future sector, these findings provided evidence for the conclusion that Brazilian Portuguese is well-advanced with regard to the process of variation and change that has been attested in the future sector since the first days of Romance languages.

## REFERENCE

---

- Abrams, M. 1971. *A Glossary of Literary Terms*. Holt, Rinehart and Winston: USA.
- Almeida, N. 1985. *Gramática Latina*. 20th Ed. São Paulo: Saraiva.
- Amaral, A. 1920. *O Dialecto Caipira*. São Paulo: Casa Editora "O Livro".
- Arruda, L. 2000. *Gramática de Português para Estrangeiros*. Porto: Porto Editora.
- Azevedo, A. 1965. *A Capital Federal*. Rio de Janeiro: Editora Letras e Artes.
- Azevedo, J. C. 2000. *Fundamentos de Gramática do Português*. Rio de Janeiro: Jorge Zahar Editor.
- Baleeiro, M. 1988. *O Futuro do Presente do Português Culto Falado em São Paulo*. Master's Thesis. São Paulo: State University of Campinas.
- Bandeira, Manuel. 1930. Evocação do Recife. In: *Libertinagem*. Rio de Janeiro: Pongetti. <http://www.libertinagemonline.hpg.ig.com.br/nova-pagina-2.htm>.
- Barreto, M. 1944. *Últimos Estudos: Vários*. Rio de Janeiro: EPASA.
- Barreto, M. 1980. *Novos Estudos da Língua Portuguesa*. 3rd Ed. Rio de Janeiro: Instituto Nacional do Livro e Fundação de Rui Barbosa/MEC.Presença.
- Barros, J. 1540/ 1957. *Gramática da Língua Portuguesa*. 3rd. Ed. Lisboa: Sociedade Astoria Ltda.
- Barros, J. 1540/ 1971. *Gramática da Língua Portuguesa: cartinha, gramática, diálogo em louvor da nossa linguagem e diálogo da viciosa vergonha*. Lisboa: Publicações da Faculdade de Letras da Universidade de Lisboa.
- Bechara, E. 1968. *Moderna Gramática Portuguesa: Curso Médio*. 13th. Ed. São Paulo: Companhia Editora Nacional.
- Benveniste, E. 1968. Mutations of linguistic categories. In: W. Lehmann and Y. Malkiel (eds.), *Directions for Historical Linguistics: A Symposium*. Austin: University of Texas Press.
- Bentivoglio, P. and Sedano, M. 1992. Morfosintaxis. In: *El idioma español de la Venezuela actual*. Caracas: Editorial Arte S.A.
- Bolinger, D. 1977. *Meaning and Form*. London/ New York: Longman.

- Bortoni-Ricardo, S. M. (to appear) *Nós chegueµu na escola e agora? – Leituras de Sociolinguística para Professores*. Campinas: Editora Mercado das Letras.
- Bortoni-Ricardo, S. M. 1991. *Dialect Contact in Brasília*. Manuscript. University of Brasília.
- Bortoni-Ricardo, S. M. 1998. A análise do português brasileiro em três *continua*: o *continuum* rural-urbano, o *continuum* de oralidade-letramento e o *continuum* de monitoração estilística. In: S. Grosse and K. Zimmermann (eds.), *Substandard e Mudança no Português do Brasil*. Frankfurt: TFM.
- Buarque, F. 1978. *Ópera do Malandro*. Lisboa: Editora 'O Jornal'.
- Bybee, J. and Dahl, O. 1989. The Creation of Tense and Aspect Systems in the Languages of the World. *Studies in Language* 13-1. 51-103.
- Bybee, J. and Pagliuca, W. 1987. The Evolution of Future Meaning. In: A. Ramat, O. Carruba and G. Bernini (eds.), *Papers from the 7<sup>th</sup> International Conference on Historical Linguistics*. Amsterdam Studies in the Theory and History of Linguistic Science. Amsterdam/ Philadelphia: John Benjamins.
- Bybee, J.; Perkins, R. and Pagliuca, W. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: The University of Chicago Press.
- Cardoso, W. & Cunha, C. 1978. *Estilística e Gramática Histórica*. Rio de Janeiro: Tempo Brasileiro.
- Chambers, J. K. 1995. *Sociolinguistic Theory. Chapter 5*. Oxford: Blackwell.
- Chambers, J. K. 2002. Patterns of Variation including Change. In: J. K. Chambers, P. Trudgill and N. Schilling-Estes (eds), *The Handbook of Language Variation and Change*. Oxford: Blackwell.
- Cheshire, J. 1999. Spoken Standard English. In: T. Bex and R. Watts (eds.), *Standard English: The Widening Debate*. London: Routledge.
- Comrie, B. 1985. *Tense*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Corrêa, G. 1964. *O Programa de Vernáculo: Curso Ginásial*. 2nd ed. Rio de Janeiro; Livraria Francisco Alves.
- Cunha, C. & Cintra, L. 1984/ 1999. *Nova Gramática do Português Contemporâneo*. 15th. Ed. Lisboa: Edições João Sá da Costa.

- Cyrino, S. 1994. O Objeto Nulo no Português do Brasil: um Estudo Sintático-Diacrônico. Tese de Doutorado. São Paulo: University of Campinas.
- Cyrino, S. 1996. Observações sobre a mudança diacrônica no Português do Brasil: objeto nulo e clíticos. In: I. Roberts & M. A. Kato (orgs) *Português Brasileiro - uma viagem diacrônica*. São Paulo: Editora da UNICAMP.
- De Carvalho, R. 1949. *Pequena História da Literatura Brasileira*. 8th Ed. Rio de Janeiro: Briguiet Editores.
- Duarte, M. E. 1996. Do pronome nulo ao pronome pleno: a trajetória do sujeito no Português do Brasil. In: I. Roberts & M. A. Kato (orgs) *Português Brasileiro - uma viagem diacrônica*. São Paulo: Editora da UNICAMP.
- Entwistle, W. 1973. *Las lenguas de España: castellano, catalán, vasco y gallego-portugués*. Madrid: Ediciones Istmo.
- Ferraz, M. 1980. Apresentação crítica, notas, glossário e sugestões para análise literária. In: A. Silva, *Guerras do Alecrim e da Mangerona*. Lisboa: Seara Nova.
- Ferreiro, M. 1996. *Gramática Histórica Galega*. 2nd Ed. Santiago de Compostela: Edicions Raivento.
- Figueiredo, C. 1906. *Falar e Escrever – Novos Estudos Práticos da Língua Portuguesa ou Consultório Popular de Enfermidades da Linguagem*. Vol. 1. Lisboa: Livraria Clássica Editora – A.M. Teixeira & Cia.
- Figueiredo, C. 1929. *Falar e Escrever – Novos Estudos Práticos da Língua Portuguesa ou Consultório Popular de Enfermidades da Linguagem*. Vol. 3. 4th Ed. Lisboa: Livraria Clássica Editora – A.M. Teixeira & Cia.
- Figueiredo, C. 1941. *Falar e Escrever – Novos Estudos Práticos da Língua Portuguesa ou Consultório Popular de Enfermidades da Linguagem*. Vol. 2. 4th. Ed. Lisboa: Livraria Clássica Editora – A.M. Teixeira & Cia.
- Fleischman, S. 1982a. *The Future in Thought and Language. Diachronic Evidence From Romance*. Cambridge: Cambridge University Press. [Cambridge Studies in Linguistics, 36].
- França Júnior, J. 1980. A Lotação dos Bondes. In: C. Gadelha & M. Saadi, *Teatro de França Júnior*. Vol. 1. Rio de Janeiro: Ministério da Educação e Cultura, Fundação Nacional de Arte.
- França Júnior, J. 1980. Dois Proveitos em um Saco. In: C. Gadelha & M. Saadi, *Teatro de França Júnior*. Vol. 1. Rio de Janeiro: Ministério da Educação e Cultura, Fundação Nacional de Arte.

- França Júnior, J. 1980. O Defeito de Família . In: C. Gadelha & M. Saadi, *Teatro de França Júnior*. Vol. 1. Rio de Janeiro: Ministério da Educação e Cultura, Fundação Nacional de Arte.
- França Júnior, J. 1980. O Tipo Brasileiro. In: C. Gadelha & M. Saadi, *Teatro de França Júnior*. Vol. 1. Rio de Janeiro: Ministério da Educação e Cultura, Fundação Nacional de Arte.
- Gândavo, P. 1574/ 1981. *Regras que Ensinam a Maneira de Escrever e a Ortografia da Língua Portuguesa*. Edição Fac-similada. Lisboa: Biblioteca Nacional.
- Garcia, J. M. 2000. *Introdução a Teoria e Prática do Latim*. Brasília: Editora UnB.
- Gomes, D. 1990. O Bem Amado. In: A. Mercado, *Os Falsos Mitos*. Vol. 2. Rio de Janeiro: Editora Bertrand Brasil.
- Gomes, D. 1990. O Berço do Herói. In: A. Mercado, *Os Falsos Mitos*. Vol. 2. Rio de Janeiro: Editora Bertrand Brasil.
- Guarniere, G. 1961. *A Semente: peça em 3 atos*. São Paulo: Massao Ohno.
- Guy, G. 1996. Form and function in linguistic variation. In: G. Guy, C. Feagin, D. Schiffrin & J. Baugh (eds), *Towards a Social Science of Language. Vol. 1: Variation and Change in Language and Society*. Amsterdam: John Benjamins.
- Hamilton-Faria, H. 1976. *The Farces of Gil Vicente: a study in the stylistics of satire*. Spain: Playor S.A.
- Heine, B.; Claudi, U. and Hünnemeyer, F. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: The University of Chicago Press.
- Hickey, R. 2000. Salience, stigma and standard. In: L. Wright (ed.), *The development of standard English, 1300-1800: theories, descriptions, conflicts*. Cambridge: Cambridge University Press.
- Huber, J. 1933. *Gramática do Português Antigo*. Fundação Calouste Gulbenkian: Lisboa.
- Justice, A. 1701/ 1970. *A Complete Account of the Portuguese Language*. A Scholar Press Facsimile. England: The Scholar Press Limited.
- Kury, A. 1986. *Novas Lições de Análise Sintática*. São Paulo: Editora Ática.
- Kury, A. 1989. *Para Falar e Escrever Melhor o Português*. 2nd Ed. Rio de Janeiro: Nova Fronteira.
- Labaki, A. 2000. *A Boa*. São Paulo: Boitempo Editorial.

- Labov, W. 1966. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45: 715-762.
- Labov, W. 1972a. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. 1994. *Principles of Linguistic Change, Volume I: Internal Factors*. Oxford: Blackwell.
- Labov, W. 2001. *Principles of Linguistic Change, Volume II: Social Factors. Chapter 3*. Oxford: Blackwell.
- Lapa, M. 1968. *Estilística da Língua Portuguesa*. 5th Ed. Rio de Janeiro: Acadêmica.
- Leão, D. 1608/ 1983. *Ortografia e Origem da Língua Portuguesa*. Portugal: Imprensa Nacional-Casa da Moeda.
- LeBlanc, C. and Poplack, Sh. 1999. Prescription vs. praxis: Conditional usage in French hypothetical *si*-clauses. Paper presented at NWAVE XXVIII, Toronto.
- Lins, A. and Buarque de Holanda, A. 1956. *Roteiro Literário do Brasil e de Portugal – Antologia da Língua Portuguesa*. Vol II. (Brasileiros). Rio de Janeiro: Ed. Livraria José Olympio.
- Lins, A. and Buarque de Holanda, A. 1966. *Roteiro Literário de Portugal e do Brasil – Antologia da Língua Portuguesa*. Vol I. (Portugueses). Rio de Janeiro: Ed. Civilização Brasileira S. A.
- Lippi-Green, R. 1997. *English with an accent: language, ideology, and discrimination in the United States*. New York: Routledge.
- Machado Filho, A. 1938. *Escrever Certo*. 2nd Ed. Rio de Janeiro: Editora ABC.
- Maia, C. 1986. *História do Galego-Português: Estado linguístico da Galiza e do Noroeste de Portugal desde o século XII ao século XVI*. Coimbra: Fundação Calouste Gulbenkian – Gráfica de Coimbra.
- Mateus, M. H.; Brito, A.; Duarte, I. & Faria, I. 1983. *Gramática da Língua Portuguesa: Elementos para a descrição da estrutura, funcionamento e uso do português actual*. Coimbra: Almedina.
- Mattoso Câmara, J. 1967. *A Forma Verbal Portuguesa em -ria*. Georgetown: Georgetown University Press.

- Mattoso Câmara, J. 1972. *The Portuguese Language*. Chicago; The University of Chicago Press.
- Mattoso Câmara, J. 1980. *Princípios de Linguística Geral*. Rio de Janeiro: Padrão.
- Melo Bacelar, B. 1783/ 1996. *Gramática Filosófica da Língua Portuguesa*. Reprodução Fac-Similada da Edição de 1783. Lisboa: Academia Portuguesa da História.
- Melo, G. 1957. *Iniciação à Filologia Portuguesa*. 2nd Ed. Rio de Janeiro: Ao Livro Técnico.
- Milroy, J. 1999. The Consequences of Standardization in Descriptive Linguistics. In: T. Bex and R. Watts (eds.), *Standard English: The Widening Debate*. London: Routledge.
- Milroy, J. 2000. Historical description and the ideology of the standard language. In: L. Wright (ed.), *The development of standard English, 1300-1800: theories, descriptions, conflicts*. Cambridge: Cambridge University Press.
- Milroy, J. and Milroy, L. 1999. *Authority in language: Investigating Standard English. Chapters 2, 3 and 4*. London: Routledge.
- Naro, A. 1981. The Social and Structural Dimensions of a Syntactic Change. *Language*, 57 (1). New York: Language.
- Naro, A.; Görski, E. and Fernandes, E. 1999. Change without change. *Language Variation and Change*, 11. Cambridge: Cambridge University Press.
- Nascentes, A. 1965. *Idioma Nacional*. 5th ed. Livraria Acadêmica: Rio de Janeiro
- Neves, M. H. 2000. *Gramática de Usos do Português*. São Paulo: Editora UNESP.
- Nunes, J. 1910. *Compêndio de Gramática Histórica Portuguesa: Fonética e Morfologia*. 8th Ed. Lisboa: Clássica Editora.
- Oliveyra, F. 1536/ 2000. *Gramática da Linguagem Portuguesa*. In: A. Torres & C. Assunção (orgs.) Lisboa: Academia de Ciências de Lisboa.
- Pena, M. 1968. O Juiz de Paz da Roça. In: *Comédias de Martins Pena*. Rio de Janeiro: Tecnoprint Gráfica.
- Pena, M. 1968. Os Três Médicos. In: *Comédias de Martins Pena*. Rio de Janeiro: Tecnoprint Gráfica.
- Pena, M. 1972. O Noviço. In: *Comédias*. Rio de Janeiro: Agir Editora.
- Pena, M. 2000. *Os Dous ou O Inglês Maquinista*. Rio de Janeiro: Civilização Brasileira.

- Poplack, Sh. 1992. The inherent variability of the French subjunctive. In: C. Lauefer and T. Morgan (eds.), *Theoretical Studies in Romance Linguistics*. Amsterdam: Benjamins.
- Poplack, Sh. 2003. Prescrição e Práxis na Evolução da Gramática. *Talk delivered at the Annual Conference of Brazilian Linguistic Association*.
- Poplack, Sh. and Tagliamonte, S. 2000. The grammaticization of going to in (African American) English. *Language Variation and Change*, 11. Cambridge: Cambridge University Press.
- Poplack, Sh. and Tagliamonte, S. 2001. *African American English in the Diaspora. Chapters 5 and 8*. U.K.: Blackwell.
- Poplack, Sh. and Turpin, D. 1999. Does the FUTUR have a future in (Canadian) French? *Probus*, 11: 133-164.
- Poplack, Sh.; Van Herk, G. and Harvie, D. 2002. Deformed in the dialects: an alternative history of non-standard English. In: R. Watts and P. Trudgill (eds.), *Alternative Histories of English*. London: Routledge.
- Prado, D. A. 1999. *História Concisa do Teatro Brasileiro: 1570-1908*. São Paulo: Editora da Cidade de São Paulo.
- Rand, D. and Sankoff, D. 1990. GoldVarb: A variable rule application for the Macintosh. Version 2.0. Centre de Recherches Mathématiques. Canada: Université de Montréal.
- Rebello, L. F. 1989. *História do Teatro Português*. 4<sup>th</sup> Ed. Portugal: Publicações Europa-América.
- Reis Lobato, A. 1771. *Arte da Grammatica da Lingua Portugueza*. Lisboa; Regia Officina Typografica.
- Roberts, I. and Kato, M. A. (orgs.) 1996. *Português Brasileiro - uma viagem diacrônica*. São Paulo: Editora da UNICAMP.
- Rossi, P. 1945. *Portuguese: The Language of Brazil*. New York: Henry Holt and Company.
- Said Ali, M. 1921/ 2001. *Gramática Histórica da Língua Portuguesa*. Brasília: Editora Melhoramentos and Editora Universidade de Brasília.
- Said Ali, M. 1964. *Gramática Secundária e Gramática Histórica da Língua Portuguesa*. 3rd Ed. Brasília: Ed. Universidade de Brasília.
- Said Ali, M. 1969. *Gramática Secundária da Língua Portuguesa*. 8th Ed. São Paulo: Edições Melhoramentos.

- Sankoff, D. 1988a. Sociolinguistic and Syntactic Variation. In: F. Newmeyer (ed.), *Linguistics: The Cambridge Survey*. Cambridge: Cambridge University Press.
- Sankoff, D. 1988b. Variable Rules. In: U. Ammon, N. Dittmar and K. Mattheier (eds.), *Sociolinguistics: an International Handbook of the Science of Language and Society*. Vol. 2. Walter de Gruyter.
- Santos, A. 1997. *O Futuro Verbal no Português do Brasil em Variação*. Master's Thesis. Brasília: University of Brasilia.
- Sardinha, L. & Ramos, L. 2000. *Prontuário e Conjugação de Verbos*. Lisboa: Didáctica Editora.
- Scherre, M. and Naro, A. 1999. *Tópicos de interpretação do Varbrul*. Manuscript. University of Brasilia.
- Silva Junior, M. 1887. *Noções de Grammatica Portugueza*. Rio de Janeiro: J. G. de Azevedo Editor.
- Silva Neto, S. 1957/ 1986. *História da Língua Portuguesa*. Rio de Janeiro: Instituto Nacional do Livro.
- Silva, A. (O Judeu). 1957. A Vida de Esopo. In: *Duas Comédias de Antônio José, O Judeu*. Rio de Janeiro: Editora Civilização Brasileira.
- Silva, A. (O Judeu). 1980. *Guerras do Alecrim e Mangerona*. Lisboa: Seara Nova.
- Silveira Bueno, F. 1967. *A Formação Histórica da Língua Portuguesa*. 3rd Ed. São Paulo: Edição Saraiva.
- Silveira Bueno, F. 1968. *Gramática Normativa da Língua Portuguesa: Curso Superior*. 7th Ed. São Paulo: Edição Saraiva.
- Sletsjõe, L. 1965. *O Elemento Cénico em Gil Vicente*. Lisboa: Casa Portuguesa.
- Soares Barboza, J. 1852. *Compendio de Grammatica Portugueza*. 2nd Ed. Pernambuco: Typ. Dos Editores Proprietarios Santos & C.<sup>a</sup>.
- Taylor, P. 1924. 'The Construction *Habēre*-with-Infinitive in Alcuin, as an Expression of the Future'. *Romanic Review* 15: 123-7.
- Tagliamonte, S. 2002. Comparative Sociolinguistics. In: J. K. Chambers, P. Trudgill and N. Schilling-Estes (eds), *The Handbook of Language Variation and Change*. Oxford: Blackwell.
- Teatro Costumbrista Brasileiro - Martins Pena, Macedo, Alencar, França Júnior, Artur Azevedo*. 1961. Rio de Janeiro: Instituto Nacional do Livro.

- Thomas, E. 1969. *The Syntax of Spoken Brazilian Portuguese*. Nashville: Vanderbilt University Press.
- Thomas, E. 1974. *A Grammar of Spoken Brazilian Portuguese*. Nashville: Vanderbilt University Press.
- Trudgill, P. 1999. Standard English: What it isn't. In: T. Bex and R. Watts (eds.), *Standard English: The Widening Debate*. London: Routledge.
- Van Achter, E.; Monteiro, J.; Teixeira, J. & Duarte, M. 1996. *Estudar o Verbo – Exercícios Práticos para Estrangeiros*. Coimbra: Minerva Editora.
- Van Herk, G. 2002. A Message from the Past: Past Temporal Reference in Early African American Letters. Ph.D. Dissertation. Ottawa: University of Ottawa.
- Veríssimo, J. 1998. História da Literatura Brasileira: de Bento Teixeira (1601) a Machado de Assis (1908). 7th Ed. Rio de Janeiro: Topbooks Editora.
- Vicente, G. 1975. Auto da Índia. In: *Sátiras Sociais*. Edição Anotada. Publicações Europa-América.
- Vicente, G. 1975. Farsa dos Almocreves. In: *Sátiras Sociais*. Edição Anotada. Publicações Europa-América.
- Vicente, G. 1975. Romagem de Agravados. In: *Sátiras Sociais*. Edição Anotada. Publicações Europa-América.
- Vicente, G. 1998. Auto da Barca do Inferno. In: S. Spina, *Gil Vicente*. 32nd. Ed. São Paulo: Ateliê Editorial.
- Vicente, G. 1998. Farsa de Inês Pereira. In: S. Spina, *Gil Vicente*. 32nd. Ed. São Paulo: Ateliê Editorial.
- Vicente, G. 1998. O Velho da Horta. In: S. Spina, *Gil Vicente*. 32nd. Ed. São Paulo: Ateliê Editorial.
- Vieyra, A. 1809. *A New Portuguese Grammar in Four Parts*. 7th. Ed. London: F. Wingrave.
- Watts, R. 1999. The social construction of Standard English: Grammar writers as a 'discourse community'. In: T. Bex and R. Watts (eds.), *Standard English: The Widening Debate*. London: Routledge.
- Watts, R. 2000. Mythical strands in the ideology of prescriptivism. In: L. Wright (ed.), *The development of standard English, 1300-1800: theories, descriptions, conflicts*. Cambridge: Cambridge University Press.

Weinreich, U.; Labov, W. and Herzog, I. 1968. Empirical foundations for a theory of language change. In: W. Lehmann and Y. Malkiel (eds.), *Directions for Historical Linguistics: A Symposium*. Austin: University of Texas Press.

Williams, E. 1944. *An Introductory Portuguese Grammar*. 5<sup>th</sup> Ed. New York: F. S. Crofts & CO.

Williams, E. 1968. *From Latin to Portuguese: Historical Phonology and Morphology of the Portuguese Language*. 2<sup>nd</sup> Ed. Philadelphia: University of Pennsylvania Press.

## APPENDIX 1 - GRAMMARS, USAGE MANUAL, PRESCRIPTIVE LETTERS

---

- Amaral, A. 1920. *O Dialecto Caipira*. São Paulo: Casa Editora “O Livro”.
- Arruda, L. 2000. *Gramática de Português para Estrangeiros*. Porto: Porto Editora.
- Azevedo, J. C. 2000. *Fundamentos de Gramática do Português*. Rio de Janeiro: Jorge Zahar Editor.
- Barreto, M. 1944. *Últimos Estudos: Vários*. Rio de Janeiro: EPASA.
- Barreto, M. 1980. *Novos Estudos da Língua Portuguesa*. 3rd. Ed. Rio de Janeiro: Instituto Nacional do Livro e Fundação de Rui Barbosa/MEC.Presença.
- Barros, J. 1540/ 1957. *Gramática da Língua Portuguesa*. 3rd. Ed. Lisboa: Sociedade Astoria Ltda.
- Barros, J. 1540/ 1971. *Gramática da Língua Portuguesa: cartinha, gramática, diálogo em louvor da nossa linguagem e diálogo da viciosa vergonha*. Lisboa: Publicações da Faculdade de Letras da Universidade de Lisboa.
- Bechara, E. 1968. *Moderna Gramática Portuguesa: Curso Médio*. 13th. Ed. São Paulo: Companhia Editora Nacional.
- Corrêa, G. 1964. *O Programa de Vernáculo: Curso Ginásial*. 2nd ed. Rio de Janeiro; Livraria Francisco Alves.
- Cunha, C. & Cintra, L. 1984/ 1999. *Nova Gramática do Português Contemporâneo*. 15th. Ed. Lisboa: Edições João Sá da Costa.
- Figueiredo, C. 1906. *Falar e Escrever – Novos Estudos Práticos da Língua Portuguesa ou Consultório Popular de Enfermidades da Linguagem*. Vol. 1. Lisboa: Livraria Clássica Editora – A.M. Teixeira & Cia.
- Figueiredo, C. 1941. *Falar e Escrever – Novos Estudos Práticos da Língua Portuguesa ou Consultório Popular de Enfermidades da Linguagem*. Vol. 2. 4th. Ed. Lisboa: Livraria Clássica Editora – A.M. Teixeira & Cia.
- Figueiredo, C. 1929. *Falar e Escrever – Novos Estudos Práticos da Língua Portuguesa ou Consultório Popular de Enfermidades da Linguagem*. Vol. 3. 4th Ed. Lisboa: Livraria Clássica Editora – A.M. Teixeira & Cia.

- Gândavo, P. 1574/ 1981. *Regras que Ensinam a Maneira de Escrever e a Ortografia da Língua Portuguesa*. Edição Fac-similada. Lisboa: Biblioteca Nacional.
- Justice, A. 1701/ 1970. *A Complete Account of the Portuguese Language*. A Scholar Press Facsimile. England: The Scholar Press Limited.
- Kury, A. 1986. *Novas Lições de Análise Sintática*. São Paulo: Editora Ática.
- Kury, A. 1989. *Para Falar e Escrever Melhor o Português*. 2nd Ed. Rio de Janeiro: Nova Fronteira.
- Lapa, M. 1968. *Estilística da Língua Portuguesa*. 5th Ed. Rio de Janeiro: Acadêmica.
- Leão, D. 1608/ 1983. *Ortografia e Origem da Língua Portuguesa*. Portugal: Imprensa Nacional-Casa da Moeda.
- Machado Filho, A. 1938. *Escrever Certo*. 2nd Ed. Rio de Janeiro: Editora ABC.
- Mateus, M. H.; Brito, A.; Duarte, I. & Faria, I. 1983. *Gramática da Língua Portuguesa: Elementos para a descrição da estrutura, funcionamento e uso do português actual*. Coimbra: Almedina.
- Melo Bacelar, B. 1783/ 1996. *Gramática Filosófica da Língua Portuguesa*. Reprodução Fac-Similada da Edição de 1783. Lisboa: Academia Portuguesa da História.
- Neves, M. H. 2000. *Gramática de Usos do Português*. São Paulo: Editora UNESP.
- Oliveira, F. 1536/ 2000. *Gramática da Linguagem Portuguesa*. In: A. Torres & C. Assunção (orgs.) Lisboa: Academia de Ciências de Lisboa.
- Reis Lobato, A. 1771. *Arte da Grammatica da Lingua Portugueza*. Lisboa; Regia Officina Typografica.
- Rossi, P. 1945. *Portuguese: The Language of Brazil*. New York: Henry Holt and Company.
- Said Ali, M. 1964. *Gramática Secundária e Gramática Histórica da Língua Portuguesa*. 3rd Ed. Brasília: Ed. Universidade de Brasília.
- Said Ali, M. 1969. *Gramática Secundária da Língua Portuguesa*. 8th Ed. São Paulo: Edições Melhoramentos.
- Sardinha, L. & Ramos, L. 2000. *Prontuário e Conjugação de Verbos*. Lisboa: Didáctica Editora.
- Silva Junior, M. 1887. *Noções de Grammatica Portugueza*. Rio de Janeiro: J. G. de Azevedo Editor.

- Silveira Bueno, F. 1968. *Gramática Normativa da Língua Portuguesa: Curso Superior*. 7th Ed. São Paulo: Edição Saraiva.
- Soares Barboza, J. 1852. *Compendio de Grammatica Portugueza*. 2nd Ed. Pernambuco: Typ. Dos Editores Proprietarios Santos & C.<sup>a</sup>.
- Thomas, E. 1969. *The Syntax of Spoken Brazilian Portuguese*. Nashville: Vanderbilt University Press.
- Thomas, E. 1974. *A Grammar of Spoken Brazilian Portuguese*. Nashville: Vanderbilt University Press.
- Van Achter, E.; Monteiro, J.; Teixeira, J. & Duarte, M. 1996. *Estudar o Verbo – Exercícios Práticos para Estrangeiros*. Coimbra: Minerva Editora.
- Vieyra, A. 1809. *A New Portuguese Grammar in Four Parts*. 7th. Ed. London: F. Wingrave.
- Williams, E. 1944. *An Introductory Portuguese Grammar*. 5<sup>th</sup> Ed. New York: F. S. Crofts & CO.

## APPENDIX 2 - PLAYS

---

- Azevedo, A. 1965. *A Capital Federal*. Rio de Janeiro: Editora Letras e Artes.
- Buarque, F. 1978. *Ópera do Malandro*. Lisboa: Editora 'O Jornal'.
- França Júnior, J. 1980. Dois Proveitos em um Saco. In: C. Gadelha & M. Saadi, *Teatro de França Júnior*. Vol. 1. Rio de Janeiro: Ministério da Educação e Cultura, Fundação Nacional de Arte.
- França Júnior, J. 1980. O Defeito de Família . In: C. Gadelha & M. Saadi, *Teatro de França Júnior*. Vol. 1. Rio de Janeiro: Ministério da Educação e Cultura, Fundação Nacional de Arte.
- França Júnior, J. 1980. O Tipo Brasileiro. In: C. Gadelha & M. Saadi, *Teatro de França Júnior*. Vol. 1. Rio de Janeiro: Ministério da Educação e Cultura, Fundação Nacional de Arte.
- França Júnior, J. 1980. A Lotação dos Bondes. In: C. Gadelha & M. Saadi, *Teatro de França Júnior*. Vol. 1. Rio de Janeiro: Ministério da Educação e Cultura, Fundação Nacional de Arte.
- Guarniere, G. 1961. *A Semente*: peça em 3 atos. São Paulo: Massao Ohno.
- Gomes, D. 1990. O Berço do Herói. In: A. Mercado, *Os Falsos Mitos*. Vol. 2. Rio de Janeiro: Editora Bertrand Brasil.
- Gomes, D. 1990. O Bem Amado. In: A. Mercado, *Os Falsos Mitos*. Vol. 2. Rio de Janeiro: Editora Bertrand Brasil.
- Labaki, A. 2000. *A Boa*. São Paulo: Boitempo Editorial.
- Pena, M. 1968. O Juiz de Paz da Roça. In: *Comédias de Martins Pena*. Rio de Janeiro: Tecnoprint Gráfica.
- Pena, M. 1972. O Noviço. In: *Comédias*. Rio de Janeiro: Agir Editora.
- Pena, M. 1968. Os Três Médicos. In: *Comédias de Martins Pena*. Rio de Janeiro: Tecnoprint Gráfica.
- Pena, M. 2000. *Os Dous ou O Inglês Maquinista*. Rio de Janeiro: Civilização Brasileira.

- Silva, A. (O Judeu). 1957. A Vida de Esopo. In: *Duas Comédias de Antônio José, O Judeu*. Rio de Janeiro: Editora Civilização Brasileira.
- Silva, A. (O Judeu). 1980. *Guerras do Alecrim e Mangerona*. Lisboa: Seara Nova.
- Vicente, G. 1975. Auto da Índia. In: *Sátiras Sociais*. Edição Anotada. Publicações Europa-América.
- Vicente, G. 1975. Farsa dos Almocreves. In: *Sátiras Sociais*. Edição Anotada. Publicações Europa-América.
- Vicente, G. 1975. Romagem de Agravados. In: *Sátiras Sociais*. Edição Anotada. Publicações Europa-América.
- Vicente, G. 1998. Farsa de Inês Pereira. In: S. Spina, *Gil Vicente*. 32nd. Ed. São Paulo: Ateliê Editorial.
- Vicente, G. 1998. Auto da Barca do Inferno. In: S. Spina, *Gil Vicente*. 32nd. Ed. São Paulo: Ateliê Editorial.
- Vicente, G. 1998. O Velho da Horta. In: S. Spina, *Gil Vicente*. 32nd. Ed. São Paulo: Ateliê Editorial.