

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PÓS-GRADUAÇÃO EM LETRAS/INGLÊS E LITERATURA CORRESPONDENTE

COMPUTER-ASSISTED ERROR ANALYSIS:
A STUDY OF PREPOSITIONAL ERRORS IN THE BRAZILIAN SUBCOMPONENT OF
THE INTERNATIONAL CORPUS OF LEARNER ENGLISH (Br-ICLE)

por

LEONARDO JULIANO RECKSI

Dissertação submetida à Universidade Federal de Santa Catarina em cumprimento parcial dos
requisitos para a obtenção do grau de

MESTRE EM LETRAS

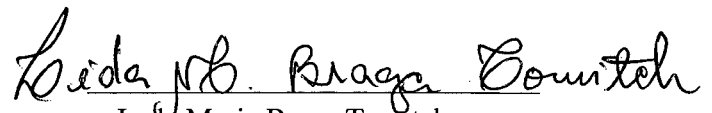
FLORIANÓPOLIS

Fevereiro 2002

Esta dissertação de Leonardo Juliano Recski, intitulada Computer-assisted Error Analysis: A study of prepositional errors in the Brazilian subcomponent of the International Corpus of Learner English (Br-ICLE), foi julgada adequada e aprovada em sua forma final, pelo Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente, da Universidade Federal de Santa Catarina, para fins de obtenção de

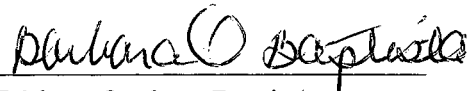
MESTRE EM LETRAS

Área de concentração: Inglês e Literatura Correspondente
Opção: Língua Inglesa e Lingüística Aplicada




Leida Maria Braga Tomitch
Coordenadora

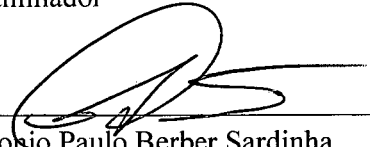
BANCA EXAMINADORA:



Bárbara Oughton Baptista
Orientadora e Presidente



Marco Antonio Esteves da Rocha
Examinador



Antonio Paulo Berber Sardinha
Examinador

Florianópolis, 18 de fevereiro de 2002.

ACKNOWLEDGEMENTS

I am indebted with all the people who in one way or another have helped me accomplish this very important step of my life.

I would like to thank Maria Lucia Vasconcelos for opening the MA doors for me when I was trying to become a special student in the program, all the PGI teachers I have had the pleasure to work with or get to know better, and specially José Luiz Meurer for helping me collect the data for this study.

I would sincerely like to express my gratitude and admiration to my adviser, Barbara Oughton Baptista, for giving me support, incentive, professional guidance and friendship in every step of this process.

Words of gratitude are also in order to Tony Berber Sardinha for valuable comments and for making available part of the data used in this dissertation and to Marco Rocha for providing valuable comments especially in relation to the statistical tests applied to the data.

I am also thankful to CAPES for providing the financial means for this journey

Thanks to my PC, for never bugging on me and king Neptune for the good swells along the way.

I am also grateful to my very good friends Márcio and Lú, Leo-bro, Gustavo and Roberta, Nilson, David, my sister Patrícia, and especially to my girlfriend Cristiana, for putting up with my occasional bad moods and for all the trust, support and friendship they have passed on to me.

Finally, no words to thank enough my greatest supporter and encourager, the one who has always believed me and given me love and strength to carry on – my *mother*. Thank you mother.

ABSTRACT

COMPUTER-ASSISTED ERROR ANALYSIS:

A STUDY OF PREPOSITIONAL ERRORS IN THE BRAZILIAN SUBCOMPONENT OF THE INTERNATIONAL CORPUS OF LEARNER ENGLISH (Br-ICLE)

LEONARDO JULIANO RECKSI

UNIVERSIDADE FEDERAL DE SANTA CATARINA
2002

Supervising Professor: Barbara Oughton Baptista

Computerized text analysis programs have been available for use on personal computers for some time. In this thesis the technique of computer-assisted error analysis, a new approach to the analysis of learners' errors is introduced, with a focus on prepositional errors. It is suggested that traditional error analysis suffered from a number of major weaknesses and that by doing computer-assisted error analysis it is possible to give new impetus to the Error Analysis enterprise. The data used in this study consists of a 33.754-word corpus of English written by Portuguese-speaking university students of English at advanced level (the Brazilian subcomponent of the International Corpus of Learner English - Br-ICLE). *Wordsmith's Wordlist* was used to obtain the ten most frequent prepositions in the Br-ICLE. After correction by two university professors who are native speakers, the corpus was annotated for prepositional errors using a comprehensive error classification. This stage was a computer-assisted process supported by an error editor. The error-tagged corpus was analyzed using standard text retrieval software tools and it was used to generate comprehensive lists of specific prepositional error types, count and sort them in various ways and view them in their contexts and alongside instances of non-errors. Results suggest that this approach provides a

new way of discovering important prepositional patterns of learner writing, in particular areas of persistent difficulty, It is suggested that a corpus tagged for prepositional errors can improve textbooks and pedagogical tools and help teachers, which, being more centered on learners' difficulties, should become more efficient.

RESUMO

COMPUTER-ASSISTED ERROR ANALYSIS:

A STUDY OF PREPOSITIONAL ERRORS IN THE BRAZILIAN SUBCOMPONENT OF THE INTERNATIONAL CORPUS OF LEARNER ENGLISH (Br-ICLE)

LEONARDO JULIANO RECKSI

UNIVERSIDADE FEDERAL DE SANTA CATARINA
2002

Orientadora: Barbara Oughton Baptista

Programas para a análise de texto para microcomputadores já estão disponíveis há algum tempo. A técnica de análise de erros preposicionais auxiliada por computador, um novo enfoque para a análise de erros de aprendizes, é abordada nesta dissertação. Sugere-se que a análise de erros tradicional, praticada no passado, apresentava uma série de problemas e que através da análise de erros com o auxílio de computadores vários desses problemas podem ser resolvidos, o que por sua vez pode dar um novo ímpeto ao estudo de erros. Os dados coletados no presente trabalho consistem de um corpus de 33.754 palavras de inglês produzidas em redações por alunos universitários brasileiros estudantes de inglês em nível avançado (o subcomponente brasileiro do Corpus Internacional de Inglês de Aprendizes - Br-ICLE). *Wordsmith Wordlist* foi utilizado para identificar a dez preposições mais frequentes no corpus. Após a correção feita por dois professores universitários nativos da língua, o corpus foi anotado para erros preposicionais utilizando-se uma categorização de erros compreensiva. Este estágio foi um processo realizado com o computador, apoiado por um editor de erros. O corpus anotado com erros preposicionais foi manipulado com softwares padrão para a recuperação de textos, sendo possível gerar listas de específicos erros preposicionais, contá-los e separá-los de diversas formas e visualizá-los em seus contextos e junto com instâncias de

usos corretos. Os resultados sugerem que esta técnica representa uma maneira inovadora de descobrir importantes características na escrita de aprendizes com relação ao uso de preposições em áreas de persistente dificuldade. Desta maneira, sugere-se que um corpus anotado de erros preposicionais pode melhorar a qualidade de livros didáticos e matérias pedagógicas e auxiliar professores, que estando mais cientes das dificuldades dos aprendizes, provavelmente podem ser mais eficientes.

Nº de páginas: 137

Nº de palavras: 26.832

TABLE OF CONTENTS

	Page
ABSTRACT	iv
RESUMO	vi
INTRODUCTION	1
CHAPTER ONE - GENERAL REVIEW OF THE LITERATURE	4
1. Introduction	4
1.1 Contextualizing Error Analysis Historically	6
1.1.1 Contrastive Analysis	7
1.1.2 Error Analysis	10
1.1.3 Interlanguage	13
1.2 Defining Error	14
1.3 Computer-assisted Error Analysis	17
1.4 Conclusion	20
CHAPTER TWO – PREPOSITIONS	22
2. Introduction	22
2.1 Difficulties in Using the English Prepositional System	22
2.1.1 The Preposition ABOUT	28
2.1.2 The Preposition AT	29
2.1.3 The Preposition BY	31
2.1.4 The Preposition FOR	32
2.1.5 The Preposition FROM	34
2.1.6 The Preposition IN	35
2.1.7 The Preposition OF	36
2.1.8 The Preposition ON	41
2.1.9 The Preposition TO	44
2.1.10 The Preposition WITH	45
2.2 Contrasting the English and the Portuguese prepositional systems	47
2.3 Conclusion	53
CHAPTER THREE – METHOD	54
3. Introduction	54
3.1 Subjects & Data Collection	54
3.2 Identification of errors in the corpus using the WordSmith Tools	56
3.3 Error Categorization	59
3.3.1 Lexical errors (L)	61
3.3.2 Dependent Prepositions (X...PR)	61
3.3.3 Complementation errors (X...CO)	62
3.3.4 Word Redundant (WR) - overuse errors	63
3.3.5 Word Missing (WM) - underuse errors	63
3.4 Checking doubtful cases through the British National Corpus (BNC)	63
3.5 Statistical Procedures	65

3.6 Research Questions	66
3.7 Conclusion	67
CHAPTER FOUR – RESULTS AND DISCUSSION	68
4. Introduction	68
4.1 The preposition ABOUT	68
4.1.1 Correct uses of ABOUT	69
4.1.2 Incorrect uses of ABOUT	71
4.2 The Preposition AT	72
4.2.1 Correct uses of AT	73
4.2.2 Incorrect uses of AT	74
4.3 The Preposition BY	75
4.3.1 Correct uses of BY	75
4.3.2 Incorrect uses of BY	77
4.4 The Preposition FOR	80
4.4.1 Correct uses of FOR	80
4.4.2 Incorrect uses of FOR	82
4.5 The Preposition FROM	84
4.5.1 Correct uses of FROM	84
4.5.2 Incorrect uses of FROM	85
4.6 The Preposition IN	85
4.6.1 Correct uses of IN	86
4.6.2 Incorrect uses of IN	87
4.7 The Preposition OF	91
4.7.1 Correct uses of OF	92
4.7.2 Incorrect uses of OF	94
4.8 The Preposition ON	99
4.8.1 Correct uses of ON	99
4.8.2 Incorrect uses of ON	101
4.9 The Preposition TO	103
4.9.1 Correct uses of TO	104
4.9.2 Incorrect uses of TO	105
4.10 The Preposition WITH	109
4.10.1 Correct uses of WITH	109
4.10.2 Incorrect uses of WITH	110
4.11 Addressing the research questions	112
CONCLUSION	121
REFERENCES	128
APPENDIXES	132
APPENDIX A: Word List - Frequency of first 100 words	132
APPENDIX B: Print Screen shot of WordSmith Wordlist - Statistics.....	133
APPENDIX C: Sample of concordance lines of <i>at</i>	134
APPENDIX D: Sample of concordance lines of <i>to</i> tagged as a preposition using the TOSCA tagger	135
APPENDIX E: Sample of concordance lines bearing the error code (XVPR) using WordSmith Concord	136

APPENDIX F: Print screen shot of attributing an XVPR tag and the corrected form in the text using the UCLEE 137

LIST OF ABBREVIATIONS

BNC	British National Corpus
Br-ICLE	Brazilian component of the International Corpus of Learner English
CA	Contrastive Analysis
EA	Error Analysis
EFL	English as a Foreign Language
ESL	English as a Second Language
FL	Foreign Language
ICLE	International Corpus of Learner English
ID	Idiosyncratic Dialect
IL	Interlanguage
LOCNESS	Louvain Corpus of Native English Essays
LP	Errors involving Lexical Phrases
LS	Errors involving independent prepositions – Lexical Single
MT/L1	Mother Tongue
NL	Native Language
NS	Native Speaker
NNS	Non-Native Speaker
POS	Part Of Speech
SLA	Second Language Acquisition
TL/L2	Target Language
UCLEE	Université Catholique de Louvain Error Editor
UCREL	University Centre for Computer Corpus Research
XADJCO	Errors involving the complementation of adjectives
XADJPR	Errors involving prepositions dependent on adjectives
XNCO	Errors involving the complementation of nouns
XNPR	Errors involving prepositions dependent on nouns
XPRCO	Errors involving the complementation of prepositions
XVCO	Errors involving the complementation of verbs
XVPR	Errors involving prepositions dependent on verbs
WM	Word Missing (avoidance errors)
WR	Word Redundant (overuse errors)

LIST OF FIGURES AND TABLES

FIGURES	Page
Figure 1: Points of comparison for FL learning paradigms	6
Figure 2: Different sources of contact for <i>on</i>	43
Figure 3: Multiplicity of forms between the Portuguese and the English prepositional systems	53
TABLES	
Table 1: Overlapping among English and Portuguese prepositional systems	53
Table 2: The 10 most frequent prepositions in the Br-ICLE	59
Table 3: <i>About</i> : distribution of correct uses	72
Table 4: <i>About</i> : distribution of errors	74
Table 5: <i>At</i> : distribution of correct uses	76
Table 6: <i>At</i> : distribution of errors	77
Table 7: <i>By</i> : distribution of correct uses	79
Table 8: <i>By</i> : distribution of errors	80
Table 9: <i>For</i> : distribution of correct uses	83
Table 10: <i>For</i> : distribution of errors	85
Table 11: <i>In</i> : distribution of correct uses	89
Table 12: <i>In</i> : distribution of errors	91
Table 13: <i>Of</i> : distribution of correct uses	95
Table 14: <i>Of</i> : distribution of errors	97
Table 15: <i>On</i> : distribution of correct uses	103
Table 16: <i>On</i> : distribution of errors	104
Table 17: <i>To</i> : distribution of correct uses	107
Table 18: <i>To</i> : distribution of errors	108
Table 19: <i>With</i> : distribution of correct uses	113
Table 20: <i>With</i> : distribution of errors	114
Table 21: Error rate distribution	116
Table 22: Distribution of errors among various error categories	117
Table 23: <i>In</i> , <i>on</i> , and <i>at</i> : distribution of spatial and temporal errors	119
Table 24: Distribution of error categories across the two corpora	120
Table 25: Frequency of prepositions across different corpora	123
Table 26: Overuse and underuse profile	123

INTRODUCTION

The present dissertation is a study of the ten most frequent English prepositions found in the Brazilian subcomponent of the International Corpus of Learner English (henceforth BRICLE) based on the theories of Error Analysis (EA), Contrastive Analysis (CA), and Corpus Linguistics methodology.

The impetus for the present study stems from the fact that English prepositional usage is highly anomalous and frequently a matter of collocational competence and experience rather than logic, which makes prepositions hard to learn. Long after EFL students have achieved a high level of proficiency in English, they still struggle with prepositions. Moreover, prepositions are one of the most difficult items to be acquired in English, partly because of the complexity of the English prepositional system which allows one preposition to have several meanings and functions, and also because many times the English prepositional system functions differently from the Portuguese system.

Bearing this in mind, the ultimate purpose of this study is (a) to identify and categorize the main difficulties advanced students face in dealing with prepositions and the likely reasons for the occurrence of these errors in their essay writings; (b) to provide evidence of correct use so as to indicate prepositional structures or patterns that seem to have been already acquired and therefore do not represent many problems; (c) to offer a methodological framework that ensures consistency of analysis enabling researchers working independently on a range of language varieties to produce fully comparable analyses; and (d) to suggest how teachers and material designers might do better at helping students achieve a better understanding of the range of common literal and metaphorical uses prepositions can have.

This thesis comprises four chapters. The first chapter is dedicated to a general review of the theories of EA and CA. It is suggested that these two paradigms are complementary to

each other and that, in contrast to the EA practiced in the past, where analyses were carried out manually on the basis of limited corpora, today, with the advances of computers and linguistic software, researchers and teachers can store and have access to larger databases, create specific tags for specific errors, and analyze these error-tagged files using standard text retrieval software tools, thereby making it possible to count errors, retrieve lists of specific error types, view errors in context, etc., thus optimizing EA and making it a more reliable enterprise.

The second chapter provides a brief description of the major patterns of syntactic and semantic use of the ten most frequent English prepositions found in the Br-ICLE (*about, at, by, for, from, in, of, on, to, and with*), as well as a contrast between the English and the Portuguese prepositional systems in order to point out likely sources of infelicities.

In the third chapter the methodological framework that has guided this study is described. The chapter provides information about the subjects in the Br-ICLE, how the data was collected, which softwares were used in the analyses and categorization of errors, the statistical tests applied to the data, and the research questions that this study proposes to answer.

In the fourth chapter each preposition is thoroughly analyzed in terms of correct and incorrect use. The aim of this stage is to indicate structures that seem to have been already acquired and contrast them with the ones that are still causing problems. The error categorization proposed in the previous chapter is put to use and the errors found for individual prepositions are distributed within these categories. An attempt to describe the sources for individual error categories is also brought to view. In the rest of the chapter, the research questions proposed at the end of Chapter 3 are addressed, and evidence provided in the analysis is put to use in order to back up each answer.

Finally, the conclusion restates the findings of the previous chapters and the application of the theories used in this study. It is suggested that if learners have access to appropriate corpus material, there is enormous scope for their own investigation of the role that prepositions play in discourse organization and the sorts of problems that learners face when making appropriate prepositional choices.

Overall, it can be argued that at least for this learner population, the ten prepositions under scrutiny do not represent a serious impediment to successful essay writing, since out of the 2.930 prepositions analyzed only 283 errors were found. This in itself should be viewed as very positive feedback since the learners correctly employed these prepositions 90% of the time they were used. Nonetheless, the *deviances* that did show provided a valuable indication of syntactic and semantic prepositional uses that should be more carefully dealt with in the future.

CHAPTER ONE

GENERAL REVIEW OF THE LITERATURE

1. Introduction

What is supposed to be good English for the EFL learner? We know that to err is human, and that to err at speaking or writing is an inherent feature of both NSs and NNSs. It is the way that these two groups err, and the medium where they err (written or spoken) that has attracted a lot of attention from both linguists and teachers, specially regarding the errors of the latter group. The crux of the problem is that English, in FL settings, usually has to serve two main roles. It must serve as a language for international communication, and as a language for communication on the global business scenario. Now, to serve both roles, the FL English must be intelligible to NSs. Greenbaum (1988) very incisively purports the following:

If English is to retain its value as an international language it is important that the norms of written English in countries where English is a second language do not diverge too far from those of the international written standard. (Greenbaum 1988, p. 38)

The statement above suggests that learners' interlanguage should be restricted to informal usage only in the spoken medium, and that for formal written purposes a more neutral international dialect should be adopted, which is the system prevailing in Britain and the USA. However, Kasper and Kellerman (1997) bring to our attention that "relative to native speakers' linguistic competence, learners' *interlanguage* is deficient by definition" (p.5). James (1998), makes the point that an interlanguage (IL) is a natural language in its own right and that it should be respected as such and described independently. Therefore, according to James (1998), it is considered malpractice even to compare the learners' IL with the native speakers' version in order to find the misfit – which is exactly what Error Analysis (EA) is all about. This practice, according to James (1998, p. 43), "is judging the students by

what they are not – native speakers. L2 learning research considers that learners should be judged by the standards appropriate to them, not by those used for native speakers”.

This argument can be questioned on general counts. First, it seems to be based on the misconception that we do judge learners by the same standards as we judge native speakers, whereas it is probably a natural tendency to judge nonnative speakers by less rigid standards. Furthermore, ‘something’ is always judged by a standard external and independent of that ‘something’: there is no point in judging something against itself, since this would not constitute a judgement. It is acknowledgeable that when IL users are speaking among themselves, they are using their own variety, which is a local variety. So, there would be no point in comparing this to, say, the American standard since local varieties of American English are usually nonstandard too. But when EFL students write something more ambitious than a shopping list or a note to a friend, they should write as close as possible to standard English, and specially so if they are writing for an audience beyond their locality. The point I want to make is that it is in the *written* medium that both native speakers and EFL students need to try to conform to a standard. Towards these objectives it is clear that learners can benefit from finding out what is *not* possible in the language. And what do we call what is not possible in a language? Ill-formed sentences? Infelicities? Malformations? Why not **errors**? Now, since the number of errors and types of errors in a language are infinite, there would be no point in trying to list them all for the learners to avoid. It would obviously be more manageable to disregard all but those errors that the learners themselves do make. And how do we know they do? By doing Error Analysis. That is what this thesis is about.

Today, however, EA seems to have gone out of fashion in most language quarters. What began as a useful pedagogical tool has apparently succumbed under the light of different successive linguistic paradigms. This is a state of affairs which needs to be remedied.

The battles over the value of EA in the 1970's, the contrastivist approach of the 50's and 60's, the Interlanguage veto on comparison and the criticism of the EA approach are systematized and reinterpreted where necessary in section 1.1. An attempt to define error and the problems that encircle such definition are dealt with in section 1.2. Finally, section 1.3 aims to demonstrate that recognizing the limitations of EA does not necessarily spell its death. It is proposed instead that it should be readapted in the form of computer-assisted EA, making use of computer learner corpora.

1.1 Contextualizing Error Analysis Historically

The errors produced by EFL learners have long been a cause of concern among linguists and EFL teachers. Most of them have searched for the reasons why "second language learners stop short of native-like success in a number of areas of the L2 grammar" (Towell and Hawkins 1994, p.14).

James (1998) suggests that in EFL learning there are three 'codes' or languages to be described. Figure 1 depicts the typical learning situation EFL students are faced with. At the beginning of their learning process, the learners do not have any knowledge or command of the FL, which is a far away point on their learning horizon. Gradually, they start moving towards their FL goals.

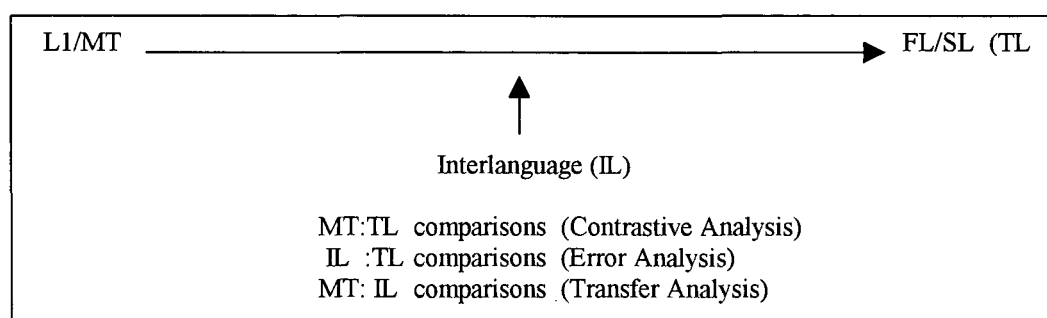


Figure 1 - Points of comparison for FL learning paradigms (James 1998, p. 3).

The language to be learned (FL/SL) can also be called the target language (TL), in the sense that the learners are actually struggling to achieve a goal, a 'target'.

The second code or language to be described is the learner's version of the TL. Teachers commonly have to make decisions about whether the learners have produced something that is right or wrong. This requires them to describe the learners' version of the TL, or as it has been called by Selinker (1972), their Interlanguage (IL), a term suggesting a halfway position between knowing and not knowing the TL. Corder (1971, p.149) prefers to call it the learner's "idiosyncratic dialect" (ID) of the TL. A label that has been used to describe the study of learners' TL is "performance analysis", which according to Corder (1975, as cited in James 1998, p. 3) is "the study of the whole performance data from individual learners", whereas the term EA is reserved for "the study of erroneous utterances produced by groups of learners" (Corder 1975, p. 207). The final code or language that needs to be described is the learner's mother tongue (MT) or L1.

When we compare these three languages (MT, IL, TL) in pairs, we have three different approaches: (a) MT / TL - Contrastive Analysis; (b) IL / TL - Error Analysis; and (c) MT / IL - Transfer Analysis. The rest of this section provides an overview of paradigms (a) CA and (b) EA, together with a discussion of how scholars have envisaged the study of IL as an alternative to both of them. The study of TA lies outside the scope of the present study.

1.1.1 Contrastive Analysis

Contrastive Analysis (CA) was the approach chosen for studying EFL learning and teaching in the 1950's and 1960's. The procedure involved first describing and contrasting features of MT and TL, and then comparing the forms and resultant meanings across two languages in order to find the misfits that would give rise to interference and error.

Even as late as 1970, Sciarone proposed that CA was valid because when one is learning a second language he/she is confronted with interference from the native language. Sciarone (1970) makes the point that the essential contribution of CA is the prediction of FL

difficulties and the description of them. In his opinion, structures of the FL that resemble those of the native language are easy, while those that differ are difficult.

However, by the late 1960's and early 1970's criticism of the CA approach began to be strongly voiced. Wardhaugh (1970), in a paper on the CA approach, observed that a decade earlier this approach was still a fairly new and exciting idea, apparently holding great promise for teaching and curriculum construction. "Now one is not so sure – and not solely as a result of the Chomskian revolution in linguistics" (p.124). Buteau (1970, p.134) added to this "The predicting power of contrastive analysis is now seriously questioned; it is being confronted with approaches that are more directly concerned with pupil performance". Carrol (1968, as cited in Bathia 1974, p. 338) too raised doubts about the reliability of the phenomenon of interference from the first language.

I have been assuming that positive and negative transfer phenomena in learning a second language are a reality. We could, in fact, ask the question whether transfer phenomena are not simply artifacts to particular training methods, or rather the absence of suitable training methods. Many examples of interference seem to be the result of what we may call *unguided imitative behavior*, or of untutored responding in terms of prior learning. (p. 97).

Dusková (1969) investigated the sources of errors made by Czech students enrolled in an English course. Her findings too suggested that students' errors were not only due to native language interference, but also to interference between the forms of the language being learned.

Buteau (1970), in her study of students' errors in the learning of French as a second language, gave further evidence of the presence of factors other than the differences between the primary system, or MT, and the secondary system, or TL. Though the writer did not state those other factors, it seems to remain true that the MT is not the only source of interference.

The general findings of this summary error-survey seem to indicate that the French sentences that correspond literally to their English equivalents are not necessarily the easiest to learn, that the probability of errors could not be assessed only from the degree of divergence of the two linguistic structures and consequently other factors of difficulty must be hypothesized. (Buteau 1970, p.139)

The practice of teaching English prepositions to Brazilian Portuguese-speaking students does not always support CA. Nevertheless, in cases where English uses the prepositions *in*, *on*, and *at* corresponding to the Portuguese equivalent *em*, we do find confusion among Brazilian learners between the three English prepositions. What does one get in English for statements like: *moro no Brazil*, *ela está na praia*, and *eles estão em casa*.

Whereas English has three different prepositions for the Portuguese sentences above, namely, *I live in Brazil*, *she is on/at the beach*, and *they are at home*, Portuguese has only one form: *em*. Cases like this do lead, however, to an interesting conclusion: when a known structure is represented in the foreign language by several structures, one can be sure to meet greater difficulties than in the reverse case. We can see then that CA, though able to predict some difficulties resulting from differences in structure, has limitations. The idea that difficulties of a foreign language can be predicted implies the supposition that corresponding structures are easy, and structures that differ, difficult. This supposition should be criticized on the grounds that is too simplistic.

According to Abbas (1995), despite the fact that CA has been the subject of innumerable criticisms for the last thirty years, it has made a significant contribution to our understanding of language teaching. He claims that CA has mainly been criticized for its overemphasis of interference errors, and for the fact it has failed to predict all errors the EFL learners are likely to commit. Nevertheless, Abbas claims that both EA and CA acknowledge the influence of the MT, and thus should be viewed as “complementary components” (Abbas 1995, p. 195), a stance also held by James (1994, p.179) who claims that EA and CA “should be viewed as complementing each other rather than as competitors”. Furthermore, Abbas contends that the use of CA and/or EA is largely determined by the type of topic being investigated. For instance, he states that the study of ‘adverbial positions’ diminishes the chances to conduct EA because this variation in position enables language learners to accommodate most

adverbials in more than one place, thus minimizing erroneous areas. Abbas (1995, p. 196) rounds off the discussion by claiming that “research on error treatment suggests that CA may be less realistic in many areas like the lexis where no reference to the practicability of CA over EA in terms of errors envisaged has been made”.

On balance, it seems to the present investigator that many of the predictions of TL learning difficulty formulated on the basis of CA turned out to be either uninformative (teachers had known about these errors already) or inaccurate: errors were predicted that did not materialize in Interlanguage, and errors did appear that the CA had not predicted. Nonetheless, it seems that ‘contrastive errors’ constitute an element substantial enough to be of interest to EA researchers. For these reasons, the CA element in EA can still be considered a very important one, and was, therefore, put to practice during the analysis of the data.

1.1.2 Error Analysis

By the late 1960’s the EA approach appeared as an alternative to the CA approach. It involved describing the learners’ interlanguage and the TL itself, followed by contrasts of the two, so as to find misfits. The novelty of EA, distinguishing it from CA, affirms James (1998), “ was that the mother tongue was not supposed to enter the picture. The claim was made that errors could be fully described in terms of the TL, without the need to refer to the L1 of learners”(p. 5).

Early work in EA was taxonomic, concentrating on the collection, description, and classification of commonly occurring errors. Corder (1967) introduced the idea that errors are a necessary part of linguistic development, and that errors are of significance because they represent the divergence between the grammar of the learner and that of the target language. He proposed that learners might have an 'inbuilt syllabus' which determines the order in which the grammar is acquired, and that studying learner errors might provide evidence to this order.

Four years later, Corder (1971) pointed towards the short life of the EA enterprise with the concept of *idiosyncratic dialect* (ID) which came as a development of the 1967 concept of *transitional competence*. It also had great resemblance to Selinker's *interlanguage* (1972) and affinity with Nemser's (1971) *approximative system*. For Corder (1971), an ID is a special sort of dialect, special in that while some of its rules are held in common with speakers of other dialects of the TL (or target dialect), these are too few to ensure interpretability of the learner's utterances by others. There are also too few individuals sharing the rules of any ID to allow us to say that it has a community of speakers. In other words, it is not a social dialect but an idiosyncratic one. It is, for example, arguable that learners of any FL who share the same MT and have been taught under similar conditions and syllabus *may* emerge speaking the same social dialect of that FL. They might encounter fewer problems understanding each other's utterances rendered in that dialect, but outsiders, including speakers of the target language i.e., native speakers, might find it unintelligible.

Corder (1971) went on to claim that there were other types of ID besides those of the FL learners: the language of the poets, child language, etc. However, he contended that FL learners were not deliberately deviant in their language like the language of the poet, so it would be wrong to refer to their repertoires as erroneous. In the case of the EFL learner, use of the label *error* would be particularly inappropriate "because it implies willful or inadvertent breach of rules, which in some sense, ought to be known" (Corder 1971, p. 152). Reference to error would only be justifiable if the rules had been deliberately flaunted or caused by performance factors, that is, if they were "cases of *failure* ... to follow a *known* rule" (Corder 1971, p. 152). Corder used two further arguments against calling the FL learner's ID erroneous. First, he objected that to do so was to "prejudge the explanation for idiosyncrasy" (p.152). We study an ID or IL in order "to discover why it is as it is, that is, to explain it and ultimately say something about the learning process" (pp. 152-53). Corder's

second argument was directed to those who called a learner's sentence ungrammatical: this is wrong, he argued, since "they are in fact *grammatical* in terms of the learner's language" (p. 155), that is in terms of the learner's ID or IL grammar.

Other, more direct criticism of EA followed Corder's. Bell (1974) referred to EA as "a recent pseudoprocedure in applied linguistics" (p. 35), and criticized EA for its poor statistical inference, the subjectivity of its interpretations of errors, and its lack of any predictive power.

Hammarberg (1974) pointed to the 'insufficiency of error analysis', which for him lied in its one-sided practice of "analyzing out the errors and neglecting the careful description of the non-errors" (p. 185). This was a narrow view, he argued, because it kept from teachers exactly the information they could put to good use, information about potential errors that learners somehow manage to avoid committing. Knowing how learners avoid certain likely errors is the first step to discovering how to help the same learners avoid the errors they fail to avoid.

Schachter (1974) also discovered what she saw as a fundamental flaw in EA – a failure to recognize that learners have a tendency to avoid TL items they are unsure about, and so not to commit errors which they would be expected to commit.

In a later article, Schachter and Celce-Murcia (1977, p. 442) commented: "EA currently appears to be the darling of the 70's", implying that the pendulum had swung too far in favor of EA and that it was time to expose some of its weaknesses. A sort of competition went on to establish the supremacy of CA over EA or the converse.

By the mid-1970's, so it seemed, EA was considered obsolete. Problems of sampling, failure to acknowledge the subjectivity of acceptability judgements and error classifications, the isolation of errors from their context, dubious statistics, the tendency to claim too broadly on the significance of the results of the analyses, and the taxonomy behind EA, were all being

bombastically condemned by many scholars, as a means to clear the theoretical ground for the new IL approach.

1.1.3 Interlanguage

The term *interlanguage* was created by Larry Selinker, in recognition of the fact that EFL learners construct a linguistic system that draws, in part, on the learner's MT but is also different from it and also from the TL. A learner's interlanguage is, therefore according to Selinker (1972), a unique system.

Nonetheless, James (1998, p. 6) points out that the distinctiveness of the interlanguage approach, brought to the field by Selinker (1972), lay “in its insistence on being wholly descriptive and eschewing comparison”. Bley-Vroman (1983, p. 15), claims that IL tries to avoid “the comparative fallacy in FL learner research”, that is “the mistake of studying the systemic character of one language by comparing it to another”.

In one sense, however, according to James (1998), interlanguage research is inescapably comparative. He conceptualizes IL in two ways. First, “it can refer to the abstraction of learner language, the aggregate of forms, processes and strategies that learners resort to in the course of tackling an additional language” (p. 7). Alternatively, James (1998) asserts that IL “can be used to refer to any one of a number of concretizations of the underlying system” (p.7). These, concretizations, following James, are sequenced in time:

IL1 develops after 100 hours of exposure, IL2 after 200 hours, and so on. The SLA researcher who studies IL developmentally or longitudinally, like the historical linguist, will be forced to comparisons of these successive stages. (James 1998, p. 7)¹.

Hawkings (1987, p. 471) has no difficulty with the comparative definition of error. He contends that “the concept of error is an intrinsically relational one. A given feature of an IL is an error *only* by comparison with the corresponding TL: seen it in its own terms the IL is a

¹ This is rather a radical view of acquiring an IL, and my own experience with learning a FL leads me to reject this view.

completely well formed system". This is exactly what EA has been claimed to be: a comparison of IL and TL.

Regarding L2 acquisition, Brown (1994) asserts that in recent years researchers and teachers have come more and more to understand that FL learning is

a process of constructing a system in which learners are consciously testing hypotheses about the target language from a number of possible sources of knowledge: limited knowledge of the target language itself, knowledge about the native language, knowledge about the communicative function of language, knowledge about language in general, and knowledge about life, human beings, and the universe. The learners, in acting upon their environment, construct what to them is a legitimate system of language in its own right - a structured set of rules that for the time being provide order to the linguistic chaos that confronts them. (p. 203)

It can be argued, then, that a possible approach to analyzing interlanguage is to study the writing of learners. Production data is presumably reflective of a learner's underlying competence. It follows that the study of the writing of learners is largely the study of the errors and correct uses of learners. Correct production yields information about the actual interlanguage system of learners, i.e., information about the TL system that learners have already acquired, and learners' errors are a register of their current perspective of the TL.

Having argued that EA is still justified today despite many of its weaknesses; I shall attempt to define error and the problems that encircle such definition in the next section. An error definition for this particular study is also proposed and put to practice in the analysis of the data (Chapters 3 and 4).

1.2 Defining Error

Many are the terms used to refer to EFL students short achievements regarding their TL: ignorance, incompleteness, ungrammaticality, infelicities, deviant language, mistake, etc. Are they all the same? Do all of them refer to the same thing and can they be used interchangeably?

In a very general way, the error analyst's objective of study, is the EFL learner's ignorance of the TL. This ignorance is the failure to attain full NS-like knowledge of the TL,

and as long as this sort of ignorance exists there will be EA. Notice that I am defining error in terms of the discrepancy between IL and the NS's version, the TL.

As a result, when we try to define errors with reference to NSs, which implies comparisons between NNSs' and NSs' utterances, the comparison should be between two categories of *people*. It is for this reason that I endorse Lennon's definition of error as "a linguistic form or combination of forms which, in the same context and under similar conditions of production, would, in all likelihood, not be produced by the native speaker counterparts" (1991a, p. 32). One of the strengths of this definition is the way it bypasses the problem of semantic intention and formal intention: what the learners wanted to communicate, and the means they employed to achieve that end. Lennon's 'in the same context' makes the meaning intention a constant, so we are left with one variable – linguistic form. Lennon's idea of associating the learners with their social counterparts is also ingenious. The counterparts will be the individuals with a similar level of education, of the same age-group, etc. Nevertheless, Lennon (1991a, p. 32) makes the claim that it is very difficult to come up with *one* universal definition for EFL error. He aptly states:

It is indeed likely that no universally applicable definition of L2 error can be formulated, and what is to be counted as an error will vary according to the situation, reference group, interlocutor, mode, style, production pressure".

According to James (1998), ignorance of the TL can be expressed in terms of two categories: Grammaticality, and Acceptability. "Grammaticality is synonymous with well-formedness. It is the grammar (not you or I) who decides whether something said by a learner is grammatical" (James 1998, p. 65). Therefore, if we can point to a bit of language and say that there are no circumstances where it could be ever said or written in this way, we are dealing with ungrammaticality². This parameter seems a bit limited in scope. Although ungrammatical, sometimes sentences can be understood if we refer not to rules, but to

² It appears to me that the appeal to grammaticality is an attempt to be objective, to take decisions such as whether an extract of language is erroneous out of the human whim.

contexts, trying to contextualize the sentence under investigation. So, on balance, the learner's intention, though ungrammatical, should be taken into account, which takes us to James' second category: acceptability. For James (1998, p. 66) acceptability "is not a theoretical but a practical notion, being determined by the use or usability of the form in question", or in other words, when a sentence is not well-formed grammatically, we usually consider it unacceptable. Nonetheless, to James (1998), even though contexts raise the degree of acceptability of doubtful sentences, judgements about the grammaticality of a sentence have to be made looking at the sentence in isolation, not in context. For this reason, he explains, to decide whether something is acceptable, even when it satisfies grammaticality, is seldom clear-cut and takes some thought.

There is a point that needs to be questioned, though. The idea that one refers to the grammar when deciding on grammaticality whereas one refers to context when deciding issues of acceptability does not seem quite right. Lennon calls our attention to the fact that "Most 'erroneous forms' are, in fact, in themselves not erroneous at all, but become erroneous only in the context of the larger linguistic units in which they occur" (1991b, p. 189). Lennon's point refers to the accompanying linguistic context or *cotext* of a possible error.

It seems plausible to argue that there is a problem with the definition of 'acceptability'. The problem with this definition is that some individuals are more skilled readers than others, so on this criterion the degree of acceptability of a text would be determined not by the text, nor by the context of its sentence, but at least in part by the processing skill of the receiver. Furthermore, it raises the question of whether it takes a native speaker to decide whether something is acceptable or not. And this seems to be the case: whether a text is acceptable or not will depend on its idiomaticity, its appropriacy, its fluency, and its coherence. All these are aspects of texts which apparently one has to be a native speaker to pass judgement about.

Nor can a text be declared acceptable only on the grounds that it is intelligible, because the sense that the reader extracts from it might not correspond to that which its writer intended.

We are now in a position where we can begin to construct a definition of 'error' for the present context of investigation: **prepositions**. *The appeal to native speaker intuitions was found adequate as a criterion for prepositional error. An error was counted whenever the two native speakers who provided acceptability judgements for the present study indicated so.* It was found that the extent of prepositional errors will usually not be limited to the preposition itself: this is the unit that would have to be replaced, deleted or added to repair the error. For such errors, consequently, the correction process seems to be mechanically simple, but recognition that an error has been made may be more complex, in the sense that large amounts of context may, in some cases, be required. According to Lennon (1991b, p. 193), "a particular problem for error analysis is that the error may be embedded in units which themselves are erroneous choices". In fact, some of the errors gathered in the present study bear witness to this (e.g. **lace of pearls*), where the learner probably was attempting *pearl necklace*. This is to emphasize once more how error can become blurred at the edges, and to imply that the idea of prepositional error as necessarily locally identifiable and traceable to a particular linguistic element may be a simplistic one.

Finally, I would like to stress that this study of errors is in no way concerned with allocating *blame*. When I say that a learner's sentence, or bit of sentence is *wrong*, I am merely using a label to refer to a discrepancy between what this particular learner *tends* to say and what the native-speaking counterparts *tend* to say.

1.3 Computer-assisted Error Analysis

As mentioned elsewhere, recognizing EA's weaknesses does not necessarily mean that the EA enterprise as a whole was unjustifiable. Instead, this section will try to show that the

EA practiced in the past can be remodeled into a new approach to learners' errors which can give it a new impetus and re-establish it as an important area of study.

According to Dagneaux et al. (1998, p. 164) traditional error analysis suffered from a number of major weaknesses, among which the following five figure prominently:

- *Limitation 1*: EA was based on heterogeneous learner data;
- *Limitation 2*: EA categories were fuzzy;
- *Limitation 3*: EA could not account for phenomena such as avoidance;
- *Limitation 4*: EA was restricted to what learners could not do;
- *Limitation 5*: EA gave a static picture of L2 learning.

The two first limitations are methodological. With respect to the data, traditional EA has often proven disappointing because of a lack of rigor in the research methods employed. An examination of the existing literature in the area of EA reveals striking discrepancies in the results, usually because there have been fundamental differences in the data and sampling methods themselves (for example, studies have used subjects with different levels of language ability, performing different tasks etc.). One of the key aims of the current project was to collect data which were comparable, ensuring that variables were rigorously controlled. All of the subjects share features like age (approximately 25 years of age), learning context (EFL), medium (writing), genre (argumentative essay writing), and length (approximately 500 words). It can be argued, then, that the present study does not suffer from *Limitation 1*.

As regards error categories, Dagneaux et al. (1998) purport that traditional EA categories also suffer from a number of weaknesses: "They are often ill-defined, rest on hybrid criteria and involve a high degree of subjectivity" (p. 164). Terms such as 'grammatical errors' or 'lexical error', claim Dagneaux et al., are rarely defined, which makes the results hard to render, as several error types, prepositional errors for instance, fall somewhere in between and it is usually difficult to assert in which of the two categories they have been

counted. In the present study, only one morpho-syntactic category (e.g. *prepositions*) is the object of study in order to ensure a more rigid categorization (in this study, prepositional errors are viewed as lexico-grammatical errors). In addition, an error editor put out by the Université Catholique de Louvain was employed in order to ensure consistency of analysis. It should enable researchers working independently on a range of language varieties to produce comparable analyses. The prepositional error categories were drawn from the manual that accompanies the error editor software (see Chapter 3 - section 3.3) which provides clear guidelines to all prepositional error types. One of the major limitations of traditional EA (*Limitation 2*) clearly does not apply here.

The other three limitations have to do with the scope of EA. Dagneaux et al. (1998) claim that EA's exclusive focus on overt errors means that both non-errors, i.e. instances of correct use, and underuse of words are disregarded. For instance, a search for all XVPR (verb dependent preposition) errors brings us down to the lexico-grammatical level and reveals that *in* is the most problematic preposition. At this stage, the analyst can draw up concordances of *in* to compare correct and incorrect uses of *in* in context and thereby get a clear picture of what the learner knows and what he/she does not know and therefore needs to be taught. This shows that computer-assisted error analysis need not be guilty of *Limitation 4*: non-errors are taken into account together with errors.

Another limitation of traditional EA (*Limitation 5*) can be met if corpora representing similar learner groups at different proficiency levels are compared.

As the learner data is in machine-readable form, text retrieval software can be used to search for specific words and phrases and one might wonder whether this method might not be a good alternative to the time-consuming process of error tagging. A search for *in*, for instance, would retrieve all the instances of the word - erroneous or not - and the analyst could easily take this as a starting-point for analysis. But how about the items that are avoided?

Error tagging provides the means of tackling the problem of avoidance in learner language, something traditional EA failed to do³.

A prepositional error tagged corpus provides access to all prepositional errors of a given learner group, some expected, others totally unexpected. It makes it possible to characterize a given learner population in terms of the proportion of their major prepositional error categories. It can be used to generate comprehensive lists of specific prepositional error types, count and sort them in various ways and view them in their contexts and alongside instances of non-errors. As Dagneaux et al. (1998) so aptly put "it is a powerful technique which will help ELT materials designers produce a new generation of pedagogical tools which, being more 'learner aware', cannot fail to be more efficient" (173).

1.4 Conclusions

Contrastive Analysis, Error Analysis, and Interlanguage Analysis are helpful and can be used in an attempt to classify and explain prepositional errors. These three types of analyses are complementary since they aim to reveal learners' difficulties in EFL learning.

Contrastive Analysis, although limited, is valid because it deals with comparisons of NS and FL in order to predict difficulties. Its limitation lies in the fact that it is only able to predict errors derived from MT interference. However, students can also produce errors which do not stem from the MT and these cannot be predicted by CA. Thus, CA will only be used as part of the explanatory stage in the EA.

Error Analysis has been involved with describing the learners' IL and comparing it to the TL. After its heyday in the seventies it has gradually fallen into disfavor and is considered "old hat" in many applied linguistic circles today. Nevertheless, it is my contention that EA is still a valid endeavor. Some of its weaknesses have been pointed out and it was suggested that

³ Although error-tagged corpora may help tackle to some extent the problem of avoidance, we can never be absolutely sure about it because learners may use very roundabout ways of expressing an idea. Only by putting

it should be re-established in the form of computer-assisted error analysis. Sardinha (1999) makes the point that computers are inserted in our lives in very common environments: from bank ATMs to very sophisticated ones such as rocket launching bases. With so many applications in our everyday life, it would be naive, Sardinha advocates, to suppose that computers cannot make an effective contribution to the field of applied linguistics. Stubbs (1996) writes about the 'heuristic' power of computers for finding out new facts about language and claims that "when computer methods are used to study large corpora, they may confirm what has been suspected or known all along: but even such confirmation will provide vastly more detailed information than would otherwise be possible" (232).

Work carried so far has demonstrated the tremendous potential of computers in helping us understand more about learner language. The investigation of learner corpora and error analyses carried out with the help of the computer may well be able to achieve the spectacular results we have witnessed in lexicography and give rise to new generation of grammars, dictionaries, EFL books and language software programs developed with the difficulties of the learner in mind.

In the following chapter, I shall describe the major patterns of use of the ten most frequent prepositions (*of, in, to, for, with, on, by, about, at, and from*) according to Quirk et al. (1985), Swan (1980; 1997), Celce-Murcia and Freeman (1983; 1990), *Collings Cobuild English Guides 1 - Prepositions* (1991), Sinclair (1991), and others. First, I address EFL students' difficulties in using the English prepositional system, then, each preposition is described in terms of their major syntactic and semantic uses, finally the Portuguese and the English prepositional systems are contrasted in order to provide likely problematic areas for the subjects of this study.

EFL learners in situations where they have to use a particular structure can we be sure they have acquired it or not. Therefore, *Limitation 3*, is only hoped to be accounted for, specially because of the number of subjects (67).

CHAPTER TWO

PREPOSITIONS

2. Introduction

In studying the class of words called prepositions, one of the first difficulties for the researcher is to find a definition for the term. The other main problem is to classify them for descriptive and pedagogical purposes. Thus the aim of this chapter is to describe: (a) the difficulties in using the English prepositional system; (b) the semantic and syntactic roles/uses for each of the ten prepositions under investigation according to different grammarians; and (c) the differences between the English and the Portuguese prepositional systems in order to point out likely areas of difficulty.

2.1 Difficulties in Using the English Prepositional System

Learning to use prepositions is a very important part of the EFL student's grammatical development. Along with word order and morphology, the use of prepositions is a major device in English for indicating syntactic relations within sentences. While the relations indicated by word order and morphology are distinct and limited in number, those indicated by prepositions are many and varied (Tomasello, 1987). As a result, prepositions are notoriously difficult to learn. Long after EFL students have achieved a high level of proficiency in English, they still struggle with prepositions. Why do prepositions cause so many problems?

According to Celce-Murcia and Larsen-Freeman (1999), one of the answers to this question is that many languages use inflections to perform the roles that English prepositions do. Secondly, in their spatial meaning, prepositions do not always match up well from one language to another, or in other words, different languages may have different prepositional

inventories, and this factor may affect the comprehension and use of spatial prepositions (Grabowski and Weiss, 1996). In addition, some prepositions do not have correspondence from one language to the other, causing even the proficient students to have problems choosing which preposition they are going to use for a particular function.

Most English prepositions have several different functions (for instance, *The Collings Cobuild: English Guides 1 - Prepositions* lists twenty-three main uses of *in*), and these may correspond to several different prepositions in another language. At the same time, different prepositions can have very similar uses like *in the morning* but *on Monday morning*, and *on a bus* but *in a car*. In some expressions, English has no preposition where one may be used in another language; in other expressions, the opposite may also be true.

Swan (1997) observes that when we use verbs after prepositions, we use *-ing* forms, not infinitives, and that prepositions are sometimes dropped before conjunctions and sometimes not (e.g., *I'm not certain (of) what I'm supposed to do*).

To complicate matters a little further for the EFL student, words like *on*, *off*, *up*, and *down* can function both as prepositions and adverb particles as in *she ran up the stairs* (preposition), and *she rang me up yesterday* (adverb particle). Many verbs and particles are regularly used together. These combinations are rather like two-word verbs. They are often called "phrasal verbs" in grammars, and their meaning is sometimes very different from the meanings of the two parts taken separately (e.g., *look after* is not the same as *look + after*).

Concerning the acquisition of prepositions, there is general agreement (Tomasello, 1987; Walkins & Rice, 1991; Rastall, 1994; Vandeloise, 1991; Johnston & Slobin, 1979; Todaka, 1996) that spatial prepositions such as *on the table* or *in the house* are the first ones to be acquired by both L1 and L2 speakers. Tomasello (1987), reporting on the first language acquisition process of his one year-old son, provides evidence that the spatial oppositions *up-down*, *on-off*, *in-out*, and *over-under* were the first to be learned. The prepositions *with*, *by*, *to*,

for, *at*, and *of*, according to Tomasello, were learned later and were omitted and misused much more than the spatial oppositions. The likely reason for this, claims Tomasello, is that a word will be learned later and will be more frequently misused if it is polysemous (has more than one meaning). Tomasello (1987) also reports two other similar studies (Grimm 1975, Voster 1984) which have found that spatial prepositions were also learned first.

However, prepositions are not found only in reference to spatial positions and as we look at those other uses, the selection of prepositions appears arbitrary and anomalous. Todaka (1996) states that as a result of this arbitrariness, the learning of prepositions involves considerable costs in memorization and storage of information. One may be *arrested for* a crime, *accused of* it and *charged with* it. We *pay attention to* something but *take notice of* it. The student must learn *insistence on*, *respect for*, *relief from*, *good at*, etc. Rastall (1994) states that in cases as the ones quoted above, where the choice of prepositions is reduced to one, the prepositions are "merely dummy grammatical forms" (p. 229) and that as a consequence there is correspondingly an arbitrary fixing of the particular preposition by the ESL/EFL student.

Regarding the positioning of prepositions, Swan (1997, p. 440) draws to our attention the fact that in English, prepositions can come at the end of clauses in certain structures such as *wh*-questions: *Who's the present for*; in relative structures: *It's Joe that I'm really angry with*; with passives: *She likes to be looked at*; and in infinitive structures: *The village is pleasant to live in*.

There are over 100 prepositions in English. This is a very small number compared with the enormous number of verbs, nouns, and adjectives which English has. Most sentences that people produce have at least one preposition; indeed, three out of the ten most frequent words in English are prepositions: *of*, *to*, and *in*. This means that the number of times the EFL learner needs to use a particular preposition is much higher than for a lexical word such as a

noun, adjective, or verb. This means that prepositions play an important role in the structure of the English language and that they may be the cause of many difficulties for these students. Most grammarians recognize that it is difficult to systematize prepositions. Indeed, very little guidance is provided in grammar books as to which preposition is the right one to use because the reasons are complex. Moreover, it is very difficult to learn how to use prepositions correctly because most of them have several functions for the same form and different prepositions can have very similar uses. For this reason, it is very common for the FL learner to commit errors while attempting to use prepositions.

Prepositions are used as the first word in a prepositional group, which provides information about place or time, or, in a more abstract way, about relationships between people and things. A crucial problem for the EFL student is that in order to produce acceptable English, he/she needs to be able to select the right preposition. Sometimes the preposition is associated with a verb: *you need to **focus on** the most important issues*; sometimes it is associated with an adjective: *she is not very **good at** mathematics*; and in other cases it can be associated with a noun: *he is a **specialist in** arts*. Like transitive verbs, prepositions take an object, called a prepositional object. The object is normally a noun phrase. The noun phrase can be simply one word: *they spoke to **me***, or it can be a complex noun phrase: *you should be worried about **the problems of the third world countries***. The object can also be a clause built around the '-ing' form of a verb. In these cases, the '-ing' clause acts like a noun phrase: *she is very good at **making up excuses***. When the object is a personal pronoun, the object form of the pronoun must be used: *Please, don't take the blame for **him***.

Quirk et al. (1985, p. 656) assert that prepositions cannot have any of the following as a complement:

- (i) a *that*-clause (e.g., **He was surprised at that she noticed him*);

- (ii) an infinitive clause (e.g., **He was surprised at to see her*);
- (iii) a subjective case form of a personal pronoun (e.g., **He was surprised at she*);

In addition, Quirk et al. (1985, p. 657) postulate that prepositional phrases have the following syntactic functions:

- (i) POSTMODIFIER in a noun phrase - (e.g. The people *on the bus* were singing.)
- (ii) ADVERBIAL
 - (a) Adjunct (e.g., The people were singing *on the bus*.)
 - (b) Subjunct (e.g., *From a personal point of view*, I find this a good alternative.)
 - (c) Conjunct (e.g., *On the other hand*, he made no attempt to help her.)
- (iii) COMPLEMENTATION
 - (a) Complementation of a verb (e.g., We were looking *at his awful paintings*.)
 - (b) Complementation of an adjective (e.g., I'm sorry *for his parents*.)

Each of the above syntactic functions specifies where in a sentence the prepositions will typically occur. The preposition is often part of an adjunct, which means that it tends to come after a verb. If the verb is intransitive, then the preposition is likely to be the next word: *It belongs to him*. Sometimes the structure of the sentence involves putting the prepositional object in front of a verb, for example, if we want to emphasize the object: *He is difficult to deal with*; or when we are using a verb in the passive: *People like to be talked to* (Quirk et al., 1985).

When we use a relative clause, there are two possible positions for the preposition. It can come at the end of the clause: *She was the one I spoke to*; or alternatively, it can come in front of the relative clause: *She was the one to whom I talked*. Notice that putting the preposition in front of the relative pronoun makes it very formal (Quirk et al., 1985)

Prepositions also come after the linking verb *be*: *He's from Florianópolis*; or other link verbs such as *seem* or *appear*. Prepositions may be used after a noun phrase to introduce

information about the noun rather than the action described by the verb: *They had received an invitation to Bob's wedding*; where the preposition tells us more about the invitation than about the fact that they had received it. Finally, some prepositions give more precise information about the adjective and what it relates to: *We are worried about the results of the exams* (Quirk et al. 1985).

Quirk et al. (1985) point out that prepositions may also be used in two and three-word sequences (which they call complex prepositions). In two-word sequences, the first word is usually an adverb, adjective, or conjunction, and the second word a simple preposition (usually *for*, *from*, *of*, *to*, or *with*). For example:

Except for Barbara, everybody wanted to go surfing;

She had to study hard because of the final exams;

They live next to my house.

Regarding three-word prepositions, Quirk et al. (1985) assert that they usually consist of the following structure: Prep 1 + Noun + Prep 2. Quirk et al. (1985, p. 670) claim that these prepositions may be subdivided according to which prepositions function as Prep 1 and Prep 2, e.g.:

- *in + noun + of - in charge of, in front of, in view of, in spite of*
- *in + noun + with - in comparison with, in contact with, in accordance with, in common with*
- *by + noun + of - by virtue of, by means of, by way of*
- *on + noun + of - on account of, on behalf of, on top of, on (the) ground(s) of*
- Other types - *in addition to, with/in respect to, in return for, as far as, etc.*

In the survey of prepositional meanings and the difficulties they present to EFL students, to which most of this chapter is devoted, space and time will be dealt with first, and

will be followed by a more superficial exemplification of other semantic relations such as cause, goal, source, origin, instrument, etc.

2.1.1 The Preposition ABOUT

The preposition *about* meaning surrounding, it is not usually employed by American speakers who usually use the preposition *around*: *He put his arms about her*. In general, British speakers tend to use *about* and *round* where American speakers use *around*.

Another interesting aspect of *about* regards the distinction in the use of this preposition and the preposition *on* in reference to subject matter. Swan (1981, p. 16) inquires what the difference would be if we said that a book, article, lecture, etc. is *about* or *on* Africa. According to him, *on* is used to suggest that the source is a serious one. *About*, according to Swan, is usual when the information given is more general, or the style of communication is more casual. Quirk et al. (1985) share the same view and allege that *on* is chiefly reserved for formal communication (public speaking, lecturing, writing, etc.), and is therefore inappropriate for verbs like *chat* or *quarrel*. Thus *he spoke on inflation* would suggest that he was making a formal speech (gave a lecture *on*), whereas *he spoke about inflation* could refer to an informal conversation or casual allusion.

The preposition *about* may also be used with the infinitive. When we are *about to do something*, it means we are going to do it very soon: *We were about to leave when they arrived*.

About can also be found in suggestions and concerns preceded by the interrogative pronouns *how* and *what*, as in *What about going to the movies tonight?*, or in *What about Jack?* In spoken English, *about* can be used to introduce a topic you want to discuss: *Now, about your exams results David, they are not very good are they?* Celce-Murcia and Larsen-Freeman (1999) bring to our attention that *about* can also be used for time approximations:

The train leaves in about an hour; and to express the approximate degree of something: *It's about a dollar*; *It's about 40°C outside!*

2.1.2 The Preposition AT

When we regard a place as a point, without any real size, we use *at*. It may be argued that the size of this place is not important. A person who comes from the countryside will probably say that he lives *in Chapecó*, but somebody who is going by bus from Florianópolis to Rio de Janeiro will probably say that the bus stops *at Curitiba*. For the first person, Chapecó is well known and important; it has streets, houses, shops, bars, etc., but for the traveler Curitiba is just a point on a journey (Swan, 1981).

We often use *at* with the name of a place when we are interested in the activity that happens there, and not in the exact shape or dimensions of the place. For instance, if we agree to meet someone *at the airport*, we are not interested in the fact that the airport has an inside and an outside; we forget the three dimensions, and just think of the airport as a meeting place. If one says that his neighbor works *at Lojas Americanas*, he simply wants to say who his neighbor's employer is, or where his place of work is; the nature and size of the building are not important. For this reason, *at* is very often used when we talk about places of entertainment, cafés, restaurant, and about the place where people work or study (Swan, 1981).

As regards time, the preposition *at* will always be used to indicate the very specific time something is happening. For example, if we refer to somebody's death we may use three different prepositions (Celce-Murcia & Larsen-Freeman, 1983):

He died at 7:00 p.m.	}	Progressively more general
He died on a Tuesday.		
He died on January 17, 1997.		
He died in the morning.		
He died in January.		
He died in 1997.		
He died in the 90's.		

Following Celce-Murcia and Larsen-Freeman (1983, p. 262), if we are to establish a gradient of specificity between *at*, *on*, and *in* we should construct something as follows:

-----	AT	-----	ON	-----	-----	-----	IN	-----
Most								Most
Specific								General

At can also be used to express 'intended goal' or 'target' in sentences like *she smiled at me*, or *a dog snapped at his leg*. Quirk et al. (1985) draw to our attention that in certain contexts the idea of not attaining the 'intended goal' can also be expressed as in *she shot at him*, which indicates that an attempt was made (she missed him), whereas in *she shot him*, her 'intended goal' was accomplished. In other cases, where the verb is intransitive, *to* must be used if the attainment of the goal is to be stressed, e.g., *she ran at me* (denotes hostility); *she ran to me* (denotes movement towards). Quirk et al. (1985) also point out that there is a comparable difference between *at* and *to* when combined with verbs of speaking such as *roar*, *shout*, *mutter*, etc. Thus in the sentences *he shouted at me* and *he shouted to me*, the first suggests that one is being treated merely as a target, while the second implies that the shouter is communicating with me, i.e., that I am the recipient of the message.

A more detailed analysis of the errors and semantic uses involving the preposition *at* in the present corpus is undertaken in Chapter 4 - Results and Discussion.

2.1.3 The Preposition BY

Like many other prepositions, *by* appears to have many semantic uses. *By* is used when we talk about an action, when we say what we do to get the result we want (means to achieve a goal) e.g., *I killed the fly by hitting it*. *By* is also used to refer to means of transport (*by bus, by car, by train, etc*), and alternatively *by* can also mean *at the side of*; something that is *by* you is close to you.

Another meaning of *by*, referring to time, is *no latter than*. For instance, *by three o'clock* means *at or before three*, but not after. *By the time* is used with a verb, to mean not later than the moment that something happens: *By the time he arrived, we were already in bed*.

In sentences like *the accident was caused by a motorcycle*, the part of the sentence introduced by *by* is called the agent. The agent in a passive sentence is the same person or thing as the subject of an active sentence. Compare: *The accident was caused by a motorcycle* - *A motorcycle caused the accident*.

According to Celce-Murcia and Larsen-Freeman (1983), a problem for many nonnative speakers is the choice of prepositions following adjectival participles related to emotive verbs (amuse, surprise, annoy, etc.). Celce-Murcia and Larsen-Freeman claim that occasionally these prepositions are idiosyncratic, and provide as an example *interested in*. The more usual prepositions, according to them, are *by, with, or at*. Sometimes two, or even all three of these prepositions can occur after an adjectival participle; however, Celce-Murcia and Larsen-Freeman point out that there are subtle differences in meaning: *We are surprised at/by Jack's behavior*.

Flowerdew (1998) found that clause relation is very commonly signaled by the preposition *by* (+ present participle or noun phrase). In her study⁵, Flowerdew (1998) found

⁵ In Flowerdew (1998) two corpora of similar size (approximately 40.000 words) were used: the expert corpus *Global Warning: The Greenpeace Report*, and the learner corpus (LC) comprised of a sub-section of the Hong

that the preposition *by* had negative and positive semantic uses. The most striking feature of the semantic environment in which the preposition *by* occurred, according to Flowerdew, was that the accompanying causative verb was frequently attenuated by mitigating markers: *emissions can probably be lowered by using alternative forms or methods...* or *...possibly facilitated by extensive irrigation development.*

A more detailed analysis of the errors and semantic uses involving the preposition *by* in the present corpus is presented in Chapter 4 - Results and Discussion.

2.1.4 The preposition FOR

One of the most common uses of the preposition *for* is to mark indirect objects. Many times, sentences with indirect objects cause problems for the EFL student because they must be able to sort out whether a given verb takes an indirect object preceded by *to*, *for*, or *of*. Verbs which take indirect objects fall into three semantic groups according to Jacobson (1966): 'eliciting' verbs such as *ask*, 'benefactive' verbs such as *make*, and 'dative' verbs such as *give*. Each group of verbs can be associated with the type of prepositional phrase that follows it. Thus, the EFL student must learn that he/she must select *for* in the case of 'benefactive' verbs (e.g., *make*, *buy*, *cook*, *prepare*, etc). However, some sentences with prepositional objects preceded by *for* such as *Chris bought the surfboard for me* may be ambiguous. There are two possible interpretations of this sentence: (a) Chris bought it for me because I did not have the time to do it, and (b) Chris bought it for me because it was my birthday and he wanted to give me a present. Note that if the indirect object occurs directly after the verb as in *Chris bought me the surfboard* the preposition is omitted and only one interpretation is possible (benefactive).

Regarding the deletion of the preposition *for*, Celse-Murcia and Larsen-Freemant (1983) point out that *for* can be omitted (a) when it expresses a span of time: *They have worked there (for) years; (For) how long have you been working there?*; (b) in responses to questions that would cue temporal use: *How long have you surfed? (For) Ten years*; and (c) with words like *next* and *all* also indicating spans of time: *I'm going there (*for) next week; We stayed at the beach (*for) all day.*

For is used to indicate how long an action or situation lasts. It can be used to talk about the past: *She worked there for three years*; the present: *She has been working there for three years*; and the future: *By next month she will have been working there for three years*. When *for* is used to talk about a period of time continuing up to the present, it is used with the present perfect tense, as in *I've known him for a long time*. When we are talking about a particular past moment, we use *for* with the past perfect to refer to a period of time continuing up to that moment, for example: *When they arrived, we had been waiting for one hour.*

EFL students often confuse the prepositions *for* and *during*. *During* is used to indicate when something happened: *There was a storm during the night*; whereas *for* is used to indicate how long something lasted: *He was in Peru for three days*.

Swan (1980) brings to our attention the fact that *for* can be used to give the reason for an action or situation. According to him, *for* suggests that the reason is given as an afterthought, and that for this reason *for*-clauses never come at the beginning of sentences, as in *I decided to have lunch - for I was feeling hungry*.

Other uses of *for* involve (a) purpose: *They'll do anything for money*; and (b) intended destination: *He left for New York*. Phrases of purpose or destination occur as postmodifiers: *The gym for the kids is great*; as adjuncts: *They came for the party*; and as complements in copular clauses: *This book is for you* (Quirk et al. 1985, p. 696).

A thorough description of the errors involving the preposition *for* and their likely sources together with the instances where students managed to correctly employ it is given in Chapter 4 - Results and Discussion.

2.1.5 The Preposition FROM

The preposition *from* can be used if we want to specify our viewpoint regarding something or someone as in *He lives across the road from me*. It can also be used in combination with the preposition *to* to indicate a starting point and its destination as in *An email from Bob to Jack*, or if we want to specify the duration of an event or action as in *We were on vacation from March to April*.

From can also be used to express either the material or the psychological cause for a happening. For instance, in the sentence *The athletes were weak from exercising all day long*, the preposition *from* indicates the psychological and biological reason why the athletes were weak.

Another very common use of *from* is to indicate the source or origin of something or someone as in *I bought the car from Barbara* (source), or as in *He comes from Florianópolis* (origin). According to Quirk et al. (1985), when we use *from* to indicate origin, the prepositional phrase can occur not only as an adjunct, for example, *He comes from Florianópolis*, but also as a complement in copular verbs, as in *I'm from Florianópolis*, or as a postmodifier as in *This is a friend of mine from Florianópolis*.

When we want to indicate a substance from which something is derived we usually use the preposition *from* as in *Surfboards are made from polyurethane and resin* (*out of* is also possible). We can also use *from...to* when we want to talk about degrees as in *the temperature ranged from 60 to 80 degrees yesterday*, or *you can find a used surfboard from R\$100 to R\$300*.

Both errors and correct uses of the preposition *from* regarding its semantic roles are dealt with in Chapter 4 - Results and Discussion. I now turn to the description of the semantic roles of the preposition *in*, which is the second most frequent preposition in the present corpus.

2.1.6 The Preposition IN

The preposition *in* is commonly employed to indicate space relations (position, destination, area, volume) usually referring to two-dimensional objects⁶ e.g., *The horses are in the field* (where the field is conceived as an two-dimensional space enclosed by either a fence or wall so it seems like a three-dimensional place), or three-dimensional object, for example, *There are four rooms in the house* (where the house is viewed as three-dimensional object surrounded on all sides). *In* is normally used for territories such as continents and countries (*in South America, in Brazil*), provinces and counties (*in British Columbia, in Cheshire*), and city districts (*in Brooklyn*) even if the areas are not enclosed, but for cities, villages etc., we may use *at* or *in* according to the point of view (Quirk et al. 1985).

Other spatial relations expressed with *in* regard parts of the body which are "softer and more hollow" (Swan 1983, p. 86) i.e., eye, mouth, ribs, stomach, for wounds, e.g., *he was hurt in the shoulder*, and for position inside the body, i.e., brain, kidneys, heart.

Metaphorical or abstract spatial relations can also be expressed by the preposition *in*. For instance, if we want to describe the state or condition of something or someone, we can do so by using *in* in sentences such as *The boys are in danger*, or *They are in difficulties at school* (where the noun is not metaphorical but the preposition is). We may also talk about membership or participation by making use of *in* as in *He's in the army right now*.

⁶ I define objects here in a broad range of senses including all material things in the world, i.e. organisms as well as physical things.

Regarding time, the preposition *in* can be used to indicate periods longer (weeks, months, seasons, years, and centuries) or shorter (parts of a day) than a day: *in the morning; in July; in 1989; in summer; in the 70's*. When we want to refer to a period of the night we use *in* as in *I woke up several times in the night* (*during* is also possible). The preposition *in* is also used to denote spans of time into the future: *They'll finish the book in three month's time* (at the end of a period of three months starting from now). In measuring forwards from a point of time in the past, only the following construction is normal: *He finished the book in three months* (in the space of three months from when he started it) (Quirk et al. 1985). *In* is also used to say how soon something will happen, and to say how long something takes to happen, for example: *Call me again in 30 minutes; I can bake a cake in half an hour*.

There are two major semantic uses for the preposition *in*: spatial and temporal. All the errors and correct semantic uses of the preposition *in* are discussed in detail in Chapter 4 - Results and Discussions. I now turn to the description of the semantic uses of the preposition *of*, by far the most frequent preposition in the present corpus.

2.1.7 The Preposition OF

Sinclair (1991) devotes an entire chapter of his book *Corpus, Concordance, Collocation* to the description of the preposition *of*. According to him, differently from the other prepositions, which usually combine with following nouns to produce prepositional phrases that function as adjuncts in clauses, *of* combines with preceding nouns "to produce elaborations of the nominal group" (p. 83). Sinclair (1991) acknowledges the fact that *of* occasionally heads a prepositional phrase which functions as an adjunct, citing as examples *...I think of the chaps on my film course...* and *...convict these people of negligence...* (p. 83); however, the author claims that there is an "overwhelming pattern of usage being in nominal groups" (p. 83), and explains that this fact must dominate any good description of *of*. Sinclair

(1991, p. 83) takes the argument even further and questions the morpho-syntactic classification of *of* as a preposition:

It may ultimately be considered distracting to regard *of* as a preposition at all [...] we are asked to believe that the word which is by far the commonest member of its class (more than double the next) is not normally used in the structure which is by far the commonest structure for the class i.e. adjuncts [...] it is not unreasonable to expect that quite a few of the very common words in language are so unlike the others that they should be considered as unique, one-member word classes

The author sets his description of *of* inside and outside nominal groups. According to him, twenty per cent of the occurrence of *of* lies outside nominal groups, the main categories being the following: (i) a constituent of various set phrases: *of course, in spite of, out of, because of, consisting of, as a matter of fact, regardless of, in need of*, (ii) following certain verb-forms: *remind, thought, smell, heard*; and (iii) following certain adjectives: *short, capable, full*.

The structure of nominal groups, as described by Sinclair, is based on a headword which is a noun. Adjectives, verbs, numerals, determiners, etc. come in front of the noun and modify its meaning in many ways. The function of *of* "is to introduce a second noun as a potential headword" (p. 85):

this kind	of	problem	}	Likely headwords
the axis	of	rotation		
the bottle	of	port		
leaves	of	trees		

The second noun being the likely headword contrasts with what would be expected in general grammars, where the structure *the N1 of N2* would be thought as having *N1* as a headword, with *N2* as a postmodifying prepositional phrase.

Sinclair (1991) states that with both conventional and less conventional measures *N2* is the likely headword as in the following examples:

both	of	them	}	Conventional measures
a couple	of	weeks		
millions	of	cars		
some	of	these		

a fraction	of	a second	} Less conventional measures
groups	of	five	
the amount	of	water	
1.300 grams	of	cholesterol	
a series	of	curves	

The identification of the headword is the first stage in describing a nominal group, claims Sinclair (1991). He conceptualizes a headword as "the only obligatory element in the group, so it should not be capable of ready omission" (p. 86), and invites the reader to try to make sense of a sentence omitting first N2, and then N1:

- a. There are many examples of local authorities who've taken ...
- b. There are many examples who've taken ... [omission of N2]
- c. There are local authorities who've taken ... [omission of N1]

In each of these cases, claims Sinclair, it is the omission of N2 that does the greatest harm to coherence, and c. is preferable to b. Therefore, it is plausible to argue, asserts Sinclair, that the headword of a nominal group is "the main reference point to the physical world" (p. 87). He then uses this criteria to exemplify N1s which specify some part of an N2, N1s which specify specialized parts of N2s, or N1s which specify components, aspects, or attributes of N2 as the following examples drawn from Sinclair (1991 pp. 87-9) attest:

the middle	of	the street	} Focus on a part
the edge	of	the teeth	
the top	of	the pillar	
a part	of	us	
the end	of	the nipple	
the evening	of	5 th August	} Focus on a specialized part
the first week	of	the war	
the interior	of	Asia	
the point	of	detonation	
the outskirts	of	Hannover	
the whole hull	of	your boat	} Focus on a component, aspect, or attribute
the cream	of	the Cambridge theater	
a list	of	the items	
a fact	of	modern life	
the sound	of	his feet	

Support is the next category addressed by Sinclair. In this category N1 is seen as offering support to N2:

the notion	of	intelligence	}	Support
the position	of	France		
an object	of	embarrassment		
various kinds	of	economical sanctions		
many examples	of	local authorities		

Nonetheless, there are many cases where neither N1 nor N2 seems to be dominant, and where the structure requires both of them. These cases are referred to by Sinclair as *double-headed nominal groups*. One minor type of double-headed nominal group includes titles of people, places, etc., where N1 names someone or something that is related to the institution named in N2:

the Duchess	of	Bedford
the new president	of	Zaire
the Garden	of	Allah

Within double-headed nominal groups, Sinclair contends that *nominalizations*⁷ are much more prominent, and that some grammarians choose to explain these structures as clauses which have been somehow transformed into *nominal groups*. According to him, we can say that a nominal group "allows for two nouns of equal status to be chosen and connected by *of*" (p. 91). Here are some examples:

the description	of	the lady
the design	of	nuclear weapons
the killing	of	civilians
an exhibition	of	his work
control	of	the company

The last type of double-headed nominal groups introduced by Sinclair associates *of* with possession. Due to equivalences like, *the surfboard of the boy*, and *the boy's surfboard*, it is

⁷ Sinclair (1991 p. 91) defines *nominalizations* as a relationship between the two nouns. In these cases the two nouns are viewed as having a verb-object or verb-subject relationships i.e., *the payment of Social Security* - which is similar to *'x' pays Social Security* - where N2 is in an 'object' relationship to N1, or *the enthusiastic collaboration of the auctioneers* - which is similar to *auctioneers collaborated enthusiastically* - where N2 is in a 'subject' relationship to N1.

common practice to say that N2 possesses N1. In fact, *of* tends to occur many times as a postmodifier in noun phrases in a function similar to that of the genitive. Nevertheless, some distinctions can be made if we analyze the following examples:

- a. the boy has self-esteem
- b. the boy having self-esteem
- c. the boy's self-esteem
- d. the self-esteem of the boy
- e. a boy of self-esteem

Both **d.** and **e.** have postmodifying *of*-phrases. They differ in that the head of **d.** (the self-esteem) is a notional object, whereas the head of **e.** (the boy) is a notional subject. In sentence **e.**, *of* is limited to the expression of abstract attributes, as in: *a teacher of great talent* (a very talented teacher), whereas in **d.** *of* refers to the possession of this attributes. However, sometimes the idea of 'having' or 'possessing' something cannot be expressed by *of*, specially if the object possessed has concrete, physical attributes. Compare the following constructions: [1] can have either *of* or *with*, but only *with* is generally accepted in [2] (Quirk et al. (1985, p. 704):

- [1] a woman $\left\{ \begin{array}{l} \textit{of} \\ \textit{with} \end{array} \right\}$ strong feelings [abstract]
- [2] a woman $\left\{ \begin{array}{l} \textit{?of} \\ \textit{with} \end{array} \right\}$ strong hands [concrete]

The huge frequency of *of* points to the fact that there is no lack of evidence for this particular preposition; in fact, in the present corpus there is far too much evidence. *Of* is approximately every thirty-sixth word - 2.77 per cent of all the words. The description offered in this section is by no means a thorough one; it is just a simple frame of how we might deal with the embarrassment of semantic uses portrayed by the preposition *of*. Nonetheless, it should be clear by now (at least I hope) that any account of the preposition *of* should

concentrate firstly, as suggested by Sinclair (1991), on the status of headwords, then, on the distinction between single and double heads, and finally on non-nominal uses of *of*.

All errors and correct semantic uses of the preposition *of* are discussed in detail in Chapter 4 - Results and Discussions.

2.1.8 The Preposition ON

As regards spatiality, the preposition *on* can be said to have topological features which are semantically defined by the notion of *contact*. This notion of contact is usually related to lines or surfaces (topology) which are usually seen as one or two-dimensional areas. For instance, in a sentence like *Our house is on that street*, the street is viewed as a line. But *on* can also denote an area, as in *There are some rocks on that road*, or *There is a painting on the wall*, where the road and the wall may be viewed as two-dimensional areas, i.e., as surfaces. It is worth noting that the line or surface with which the object is in contact may be rotated in different ways: we may say the light is *on the ceiling*, the light switch is *on the wall*, or even that the pen is *on the floor*, as illustrated in Figure 2:

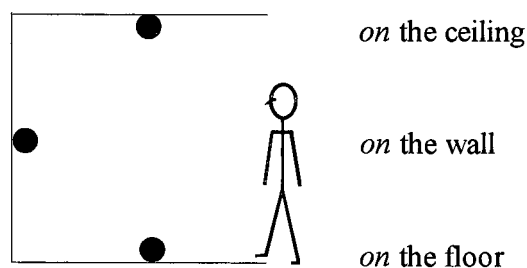


Figure 2 - Different sources of contact for *on*

It may be plausible to argue that when we refer to concrete entities we use *on* to say that something is touching or is close to a line, or something like a line (e.g., a river, a road, a frontier). We may also say that *on* can be used to indicate that something is resting on top of a surface. However, *on* can also be used for contact underneath a surface (e.g., *I've got something on the sole of my shoe*).

Although *on* indicates contact from any direction as seen in Figure 2, it can also have an array of metaphorical uses. Lindstromberg (2001 pp. 85-7) points out five such uses for *on*:

1. **The Located Object⁸ (LO) as a burden:** These are expressions which suggest that the Landmark is metaphorically burdened by the LO – (e.g., *There must be heavy taxes on cigarettes*) – a type of expression which portrays not an object (taxes) but an event (taxes on cigarettes) as the burden. Other examples are: *a weight on someone's mind, impose something on someone, shame on you, tell/inform/cheat/spy on someone, play a joke on someone, turn one's back on someone*, etc. The burden metaphor can be useful, for instance, to explain the difference between *I have something on my mind*, where the LO is a worry, and *I have something in my mind*, where the LO is something like a plan.
2. **The LO as burdensome impact:** These are expressions containing verbs or nouns which evoke burdensome impact – (e.g., *attack/assault on, war on, step on something, they turned on their teachers*).
3. **The Landmark as platform/basis or as a way:** These are expressions in which the landmark plays the role or basis for an explicit or implied action, event or state of affairs – e.g. *rely/depend/count on someone, act on the condition/assumption/premise that* etc. 'Way' expressions (e.g. *on the road to nowhere*) can also be related to basis expressions since a 'way' seems to entail a base along which someone moves.
4. **Other kinds of metaphorical contact: mental contact, visual contact:** Many types of work involve contact with something below eye level which requires prolonged attention – (e.g., *work/focus/dwell/reflect/meditate on a problem*). The difference between *wait on somebody* and *wait for somebody* also fits here, where the former entails that the landmark is present (e.g., *He is waiting on ten tables*), whereas the latter implies that the landmark is the reason for waiting, hence contact is only potential. For verbs of looking

⁸ The expressions *Located Object* (LO) and *Landmark* are glossed in accordance with their use in Cognitive Linguistics – e.g. *the cat^{LO} on the mat^{LANDMARK}* (Lindstromberg 2001, p. 79).

much the same is true – e.g. *keep an eye on*, *lay one's eye on*, *gaze on*. Nevertheless, verbs of looking collocate with other prepositions as well – e.g. *to*, *at* and *towards*. Lindstromberg (2001, p. 87) argues that the difference between *at* and *on* in the case of looking verbs seems to be that “*on* suggests visual contact which is spatially or temporarily extended and often emotionally tinged”. *At*, on the other hand, often figures in ‘target’ expressions (e.g., *throw/talk at*) since it does not define a landmark as having dimension but rather a mere point. Further, Lindstromberg (2001, p. 87) asserts that “*on* seems to resist direct collocation with verbs of impolite looking such as *stare*, *peep*, *gawk* and *gape*”.

5. **Control, influence, and effect as contact:** This refers to instances where a physical LO is in perceptible contact with a physical landmark. *On*, usually occurs in a number of control, influence, and effect expressions in which the landmark may refer to both concrete and abstract things – (e.g., *get a grip on*, *have an impact/effect/influence on*).

Other non-metaphorical spatial uses of *on* involve expressions like *on a page / on page / on the page*, and names of most parts of the body surface: *on her forehead*; *on her cheek*; *on my shoulder*; etc. *On* is also used to talk about public transport and other means of transportation such as horses, motorbikes, and bicycles(with an important exception: *in the car*). With names of streets *on* is used in American English whereas *in* is used with British English. In addition, *on* is used with the word *floor* when we want to say what part of a house or building somebody lives or works (Swan 1983).

Regarding its temporal use, *on* is normally employed to refer to days as periods of time if we mention which morning, day, etc we are thinking as in *on Friday*, *on the following day*, *on April first*, *on Christmas day*, etc., or if we describe such periods as in *on a hot afternoon*, *on that particular evening*. Quirk et al. (1985) point out that *on* is exceptionally used with complements referring to a part of a day rather than the whole day in sentences like *on Friday*

afternoon; on the following evening, etc, and that its use also extends to cases where the time segment is a part of a day which is actually mentioned: *on the evening of August 10th*.

On balance, besides spatial and time related uses, the preposition *on* seems to lend itself to uses which clearly go beyond purely spatial and temporal (i.e., metaphorical). All the errors involving *on*, as well as the instances which the subjects correctly employed it in the present corpus are extensively addressed in Chapter 4 – Results and Discussion.

2.1.8 The Preposition TO

One major use of the preposition *to* is to indicate destination. It generally accompanies verbs of motion such as *go, travel, fly, move, drive*, etc., to indicate completion of a movement in the direction of a place, as in *Barbara drove to Rio de Janeiro*. With the perfect aspect, *to* may be used interchangeably with *at* or *in* depending on the meaning we want to employ: *Leo has been to/at UFSC* (as a student), or *Leo has been to/in Florianópolis* (as a visitor). The concept of ‘implied’ motion also accounts for the use of *to*, as in *Is this the bus to Curitiba?* Very commonly *to* is used with the preposition *from* to establish a starting point/destination relationship as in *an email from me to you, he’s driving from Florianópolis to Porto Alegre*. *From...to* may also indicate duration as in *We surfed from 10 a.m. to 1 p.m. yesterday* (up to 1p.m.), but with *from* absent, only *until, till, up to* (but not always simple *to/through*) can be used as the following examples drawn from Quirk et al. (1985, p. 690) suggest:

We camped there $\left. \begin{array}{l} \textit{until} \\ \textit{till} \\ \textit{up to} \\ \textit{through} \\ \textit{*to} \end{array} \right\}$ September

Alternatively, *to* might be used to express the ‘recipient’ of an action, who is most of the times, represented by the indirect object in sentences such as *He sold/gave/lent an awesome surfboard to his friend.*

As Quirk et al. (1985) point out, *to* followed by an abstract noun of emotion (e.g. *regret, annoyance, relief, surprise, honor, delight*) can express ‘reaction’ e.g. *To my surprise, it was raining this morning.* Further, according to Quirk et al. (1985, p. 712), the reaction can also be expressed by *to* + personal pronoun or a sentence with *to* + possessive pronoun + *mind*, *in* + possessive pronoun + *opinion*, etc, to identify the person reacting:

To me,
To my mind,
In my opinion, } the waves were great.

Finally, the preposition *to* is always omitted when we use locative nouns such as *home* and *downtown* together with verbs of motion or direction e.g. *They went (*to) home, We drove (*to) downtown,* as well as when the adverbs *here* and *there* are used after the verb e.g. *They go (*to) there very often, I come (*to) here everyday* (Celce-Murcia & Larsen-Freeman 1990).

Both errors and correct uses of the preposition *to* regarding its semantic roles are dealt with in Chapter 4 - Results and Discussion.

2.1.9 The Preposition WITH

The uses of the preposition *with* were many and varied. In their chapter on prepositions, Quirk et al. (1985) refer to nine common semantic uses of *with*: spatial, pervasive, manner, means and instrument, accompaniment, support and opposition, 'having', and ingredient.

Spatial *with* can be observed in sentences like *I left the bicycle with the car* or *The wax is with the surfboard* (meaning at the same place as). The idea of pervasion, (i.e., spreading throughout), can be viewed in sentences like *The sky was filled with clouds* or *The streets*

were covered with mud. *With* expresses manner in sentences like *He was welcomed with courtesy* or *She kissed him with discretion*.

Phrases of means and instrument (which typically answer the question *How ... ?*) appear with human subjects + *with* + direct object as in *he killed her with a gun* or *they bought the house with their savings* (Quirk et al. 1985).

Perhaps one of the most common semantic uses of *with* refers to the notion of 'accompaniment'. Especially when followed by an animate complement, *with* has the meaning of 'in company with' or 'together with' as in the following examples: *John went to the concert with me*, *Richard, with many of his drunken friends, was gambling last night*. *With* is also used to express accompanying circumstances, as in *With all that wind, we find it hard to go surfing*, and to introduce a subject, as in *It all started with Jack tearing the girl's dress*. Notice that in the last two examples *with* seems to imply cause: *Because of all that wind we found it hard to go surfing* and *It all started as a result of Jack's tearing the girl's dress*.

With is also employed to convey the idea of solidarity and support, as in *The whole class was with Bob on that occasion* (= on his side), but it can also be used to convey the idea of opposition between people when it is used with verbs such as *fight*, *quarrel*, *argue*, etc, as in *He argued with his daughter last night*.

The notion of 'having' is generally expressed by *with*, especially with concrete attributes : *a boy with a yellow bicycle*, *a box with chocolates*. According to Quirk et al. (1985, p. 704) *with* can introduce a nonfinite or verbless clause as a postmodifier in a noun phrase as the examples cited above, but it can also introduce finite and verbless clauses as adverbials, as in *With so much to do, I doubt I'll have time to attend the conference*. The clausal equivalent, Quirk et al. (1985) point out, "is a participial clause expressing contingency" (p. 705): *Having so much to do, I doubt I'll have time to attend the conference*. In addition, Quirk et al. (1985) posit that since *with* introduces clauses as in *With so much to do ...*, it functions as a

subordinator, not a preposition. But how could we classify *with* in the following example: *With you I'll never feel alone*. It surely expresses the idea of accompaniment, but it can also be substituted by the participial adverbial clause *Having you, I'll never feel alone*.

Finally, with verbs of 'making', *with* indicates an ingredient, as in *the drink is made with vodka and pineapple juice*. *With* also enters in pervasive expressions such as *paved with bricks, filled with water, loaded with hay*, etc (Quirk et al. 1985, p. 711).

Both errors and correct uses of the preposition *with* regarding its semantic roles are dealt with in Chapter 4 - Results and Discussion.

In the following section, I attempt to compare both the English and the Portuguese prepositional system so as to indicate likely areas of difficulties. Further, the following section will address the issue that different languages (e.g., Portuguese and English), have different prepositional inventories and that this might be a factor which can cause problems for the subjects of this particular study.

2.2 Contrasting the English and the Portuguese prepositional systems

Examining two Portuguese grammars, Cegalla (1978) and Cunha (1978), we can notice that they deal with prepositions in the same way, i.e., in a notional approach. These grammarians merely define prepositions and try to describe their use.

According to Cunha (1978, p. 377) a preposition is "a palavra invariável que liga dois termos entre si estabelecendo que o segundo depende do primeiro, isto é, que o segundo (termo rígido) é complemento do primeiro (termo regente)". In addition, the word *termo* in the definition of prepositions does not mean that Cunha is referring to just a single word. Likewise, Cegalla (1978) defines the preposition as being "a palavra invariável que liga um termo dependente a um termo principal" (p. 175). Both definitions point out that prepositions state a relationship of dependency and subordination between two terms.

Both Cunha (1978) and Cegalla (1978) classify Portuguese prepositions into three groups: *essenciais* (essential), *acidentais* (accidental) and *locuções preposicionais* (prepositional locutions). In English we have the following classification: *simple prepositions*, which correspond to the essential and accidental ones in Portuguese; and *complex prepositions*, which correspond to prepositional locutions.

Essential prepositions are those which function strictly as prepositions: *a, ante, após, até, com, contra, de, desde, entre, para, perante*, etc. *Accidental prepositions* are those grammatical classes of words such as adjectives and adverbs which have become prepositions: *conforme, como, consoante, mediante, segundo*, etc. *Prepositional locutions* are certain expressions formed by a preposition and a noun, adverb or an adjective, plus the preposition: *abaixo de, em cima de, através de, de acordo com, por meio de*, etc.

It is also necessary to state that the syntactic relation in the sentence is determined by a fixed preposition which is selected because of its basic meaning. Thus in *concordo com você*, the verb selects the preposition *com* because of the relationship that exists between the meaning of the verb and the idea of association contained in the preposition itself. Depending on the greater or lesser intensity of meaning of the preposition, the syntactic relationship can be *fixed* (fixa), *necessary* (necessária), or *free* (livre) (Cunha 1978, p. 378).

The syntactic relation is called *fixed* when the preposition carries the meaning itself. The preposition is very important in the organization of the sentence as well as in its meaningful value. For instance, *Ninguém pode com a vida deles* (Érico Veríssimo); *Custa crer que vivem no Rio de Janeiro* (C. D. de Andrade), where *fixed* seems to mean necessary and unsubstutable.

In a *necessary relation*, the preposition links the main word to a consequent word which is syntactically necessary: *O futuro pertence a Deus*; *O homem é um grande inventor de obstáculos*, where the preposition seems to be unsubstutable.

In a *free relation*, the preposition is used but it is not absolutely necessary syntactically. Its absence does not change the meaning of the sentence as in *Encontrar com um amigo* or *Encontrar um amigo*; *Procurar por alguém*, or *Procurar alguém*.

Cegalla (1978) and Cunha (1978) argue that prepositions in Portuguese are invariable words which link two terms and whose role is to establish among these terms a relation of place, manner, time, possession, means, cause, instrument, etc. For example:

Barbara mora <i>em</i> Florianópolis.	(place)
O carro <i>do</i> Leo	(possession)
Trabalham <i>com</i> afinco	(manner)
Eles falaram <i>sobre</i> inflação	(subject matter)
Morreu <i>de</i> fome	(cause)
Surfei <i>com</i> eles	(company)

It is also worth noting that some Portuguese prepositions may also correspond to more than one relationship or function and that the same use can be expressed by different prepositions. For instance, illustrating the first case, the preposition *a* can be used to indicate:

Place: Eu vou *a* praia.

Time: Eu estudo *a* noite.

Finality: Eles foram *às* compras.

Price: Ele vendeu o barco *a* R\$ 700.00.

In the same way the preposition *em* can indicate:

Place: Ele está *em* casa.

Time: Nós chegaremos a praia *em* duas horas.

Manner: Vivemos *em* paz.

Price: A prancha foi avaliada *em* R\$ 250.00.

Finality: Vou pedi-la *em* casamento.

Comparing the English and Portuguese prepositional systems, it can be observed that the semantic roles established by grammarians to determine the function of a preposition can sometimes find correspondence in both languages. For example, the preposition *em* in

Portuguese can be used to indicate place, time, manner, price, finality, cause, etc. This preposition can relate to several forms in English - *at*, *in*, *on* - but for each one we find several of the same categories attributed to the Portuguese form: *in* can indicate place, time, manner; *at* can also denote place, time, price; and *on* may also imply place, time, and manner. Thus, there is an overlap of semantic roles between the prepositions of the two languages.

There are some prepositions in English which bear a one-to-one correspondence to those of Portuguese: *with*, *without*, *beside*, *during*, and *against*. On the other hand, one English preposition may correspond to two or more forms in Portuguese, or one form in Portuguese may correspond to more than two in English. For example:

$$On \quad \left\{ \begin{array}{l} em \\ sobre \end{array} \right\} \quad In \quad \left\{ \begin{array}{l} em \\ a \\ dentro de \end{array} \right\} \quad Dentro de \quad \left\{ \begin{array}{l} within \\ inside \\ in \end{array} \right\}$$

In examining thirty-three prepositions of both languages, one can see whether or not there is an overlap of forms and meanings. We can assume that there is no overlap of the prepositions *with*, *without*, *beside*, *during*, and *against* because they function literally, that is, there is a one-to-one correspondence in both languages, but even in these cases, there may be differences of usage in non-literal expressions such as *estar com calor* = *to feel hot*. On the other hand, we do notice an overlap of one or more forms in the use of prepositions which I have depicted in Figure 3.

Table 1 provides examples of the multiplicity of forms between the Portuguese and the English prepositional systems. As it can be noticed, Portuguese prepositions may have more than one equivalent English preposition:

Figure 3 - Multiplicity of forms between the Portuguese and the English prepositional systems

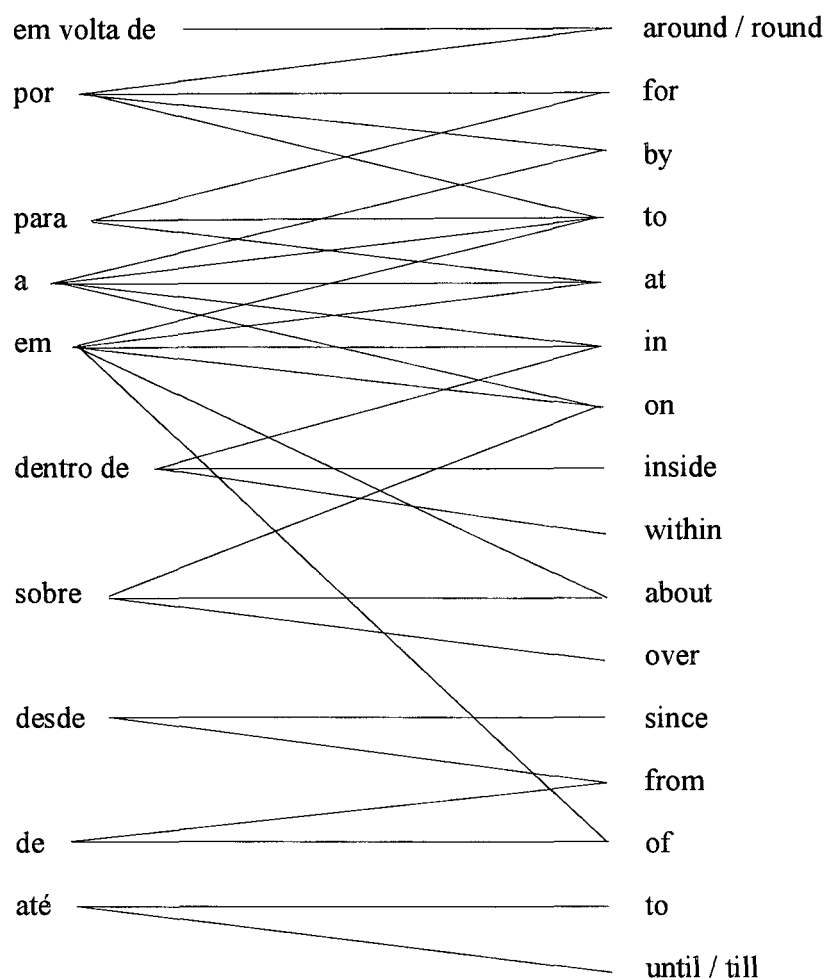


Table 1 - Overlapping among English and Portuguese prepositional systems

PREPOSITION	PORTUGUESE	ENGLISH
1. A	Ele vai à escola todos os dias	He goes to school every day
	João estuda à noite	John studies at night
	Eles chegaram a tempo	They arrived on time
	Ponha as pranchas de surf lado a lado	Put the surfboards side by side
	Ela está aqui a negócios	She's here in business
2. Até	Eu estarei ocupado até às 5	I will be busy until / till five o'clock
	Nós viajamos das 3 até às 5	We traveled from 3:00 to 5:00
3. De	Esta casa é feita de madeira	This house is made of wood
	Barbara é de Florianópolis	Barbara is from Florianópolis

4. Dentro de	O café está dentro da xícara Ela vai me pagar dentro de um mês Ele está dentro de casa	The coffee is in/inside the cup She is going to pay me within a month He is in/inside the house
5. Desde	Ele está viajando desde Sexta-feira Ele viajou desde São paulo até Curitiba	He's been traveling since Friday He traveled from São Paulo to Curitiba
6. Em	Eles estão em casa Maria mora em Curitiba O CD está na mesa Ele pensou em viajar Ela sempre sonhou em ser médica Adicione isto em sua programação	They are at home Mary lives in Curitiba The CD is on the table He thought about traveling She's always dreamed of being a doctor Add this to your programming
7. Para	Este presente é para você Eles vão para escola de manhã Olhe para mim !	This present is for you They go to school in the morning Look at me !
8. Por	Ele gosta de andar pela cidade O livro foi escrito por João Nós surfamos por duas horas Ele jurou por Deus	He likes to walk around the town The book was written by John We surfed for two hours He swore to God
9. Sobre	Eles pularam sobre o muro Nós conversamos sobre pobreza As chaves estão sobre a mesa O helicóptero voou sobre o prédio	They jumped over the wall We talked about poverty The keys are on the table The helicopter flew above/over the building

Thus, as indicated in Table 1, there is a multiplicity of forms and meanings, which in turn shows that the relations of equivalence between the two languages are very complex. Consequently, it is plausible to argue that NNSs may have many problems in using these prepositions, especially the ones which are more polysemous e.g. *em*, *para*, *por*, *dentro de*, *sobre*. This prediction is based on a comparative analysis and it is supported by evidence in the error analysis (chapter 4 - Results and Discussion).

2.3 Conclusion

The aim of this chapter was to provide an overview of the difficulties NNSs may face in attempting to use the English prepositional system. It has been shown that the semantic roles and syntactic uses of English prepositions are varied and complex. Most English prepositions (at least the ones under investigation in this study) tend to function as adjuncts, with the exception of the preposition *of* which generally functions as a noun postmodifier. By comparing the prepositional systems of English and Portuguese it was possible to conclude that only one of the ten prepositions under investigation - *with* - bears a one-to-one correspondence to Portuguese which might imply fewer difficulties for the subjects of the present study. The other prepositions - *in, on, at, by, for, to, about, from* and *of* - all present an intricate web of semantic uses, each one corresponding to more than one Portuguese preposition. Similarly, many Portuguese prepositions correspond to more than one English preposition.

In the following chapter, I present the methodological framework which has guided the present study. It addresses the following issues: data collection and subjects, softwares used to help with the error identification and categorization, NS judgments, statistical tests applied to the data, and the research questions driving the present study.

CHAPTER THREE

METHOD

3. Introduction

This chapter describes the stages or steps followed during the collection and analysis of the data. Section 3.1 provides information regarding the subjects and the way the data was collected in order to fit the standards recommended by the International Corpus of Learner English (henceforth ICLE). Section 3.2 describes how the errors were separated from correct usage and then submitted to native English university teachers for acceptability judgements. In section 3.3, both the process of error tagging the corpus using a special error editor as well as the error categories adopted in this study are described in detail. In section 3.4, the problems of coping with borderline cases of error identification are addressed. It is proposed that besides relying on native speakers' intuitions, the use of very large corpora can help the researcher find alternatives for dealing with such cases. Section 3.5 provides information regarding the types of statistical tests that were employed in the analysis of the data. Finally, the research questions which have guided the present study are addressed in section 3.6.

3.1 Subjects & Data Collection

Learner English is a very heterogeneous variety, therefore, it is essential to start from a very precise description of the population represented in the corpus. Among the variables that need to be controlled are the following: learning environment, age, mother tongue, stage of learning, and nature of the task. Comparability is one of the threads running through the current study, and thus, the learning population meets the following criteria: (a) type of learner: EFL - not ESL; (b) age: adults (by adults I mean undergraduate university students of English); (c) stage of learning: advanced (refers to undergraduate university students of

English in their third or fourth year of study); and (d) task: argumentative essay writing⁹. All relevant biographical information - such as years of English at school, prolonged stay in an English-speaking country, knowledge of other foreign languages - is encoded in a learner profile questionnaire which contributing learners were asked to fill in.

Most of the subjects were between 20-30 years of age, and wrote their essays without the help of reference tools (e.g., monolingual or bilingual dictionaries). The essays were untimed and were not held under examination conditions. Regarding the subject's previous background in English most of them had had at least 7 years of English at school, at least 2 years of English instruction at private English schools, and most of them had been studying English at university level for at least 3 years. Most of the subjects had not stayed in an English speaking country for more than 30 days. All of the subjects had Portuguese as their mother tongue and only four could speak a third language (Spanish).

The rationale behind creating a computerized corpus of learner English was to make use of advances in applied linguistics and computer technology to carry out a thorough investigation of the interlanguage (more specifically the use of prepositions) of Brazilian Portuguese English learners. The corpus was collected following the guidelines of the ICLE, which comprises corpora from different countries (18 at present), one of which is the Br-ICLE¹⁰, the subcorpus of Brazilian learners, coordinated by Tony Berber Sardinha (Pontificia Universidade Católica de São Paulo). According to Granger (1992, p 61), the ICLE project can be described as "a computerized corpus of essay writing by advanced EFL learners from various language backgrounds". When complete, the corpus will contain a minimum of

⁹ According to Granger (1996, p.18) many features of language are extremely genre-sensitive, so the type of task set will significantly alter the results obtained. Therefore, if meaningful statements are to be made about differences in usage, the type of discourse under study must be comparable.

¹⁰ The data collection for the Br-ICLE started in May 1999, and currently (July 2001) the corpus contains 33,754 running words, or 17% of the total planned, represented by 67 essays. For more information on the Br-ICLE, visit the website www.bricle.f2s.com.

200.000 words for each learner variety representing approximately 400 essays of 500 words written by at least 200 students (a student cannot contribute with more than 1000 words).

The data for the present study consists of 67 written essays produced by undergraduate students of English (24 from UFSC and 43 from PUC-SP, UFU, and UNICSUL respectively¹¹). The mean length of these texts is 504 words. The entire database amounts to a total of 33.754 tokens (words) and 4.088 types (different words).

The students were asked to write 500 words on argumentative topics that were chosen from the ones suggested in the ICLE project. Here are some of the topics students were given to choose from:

- (1) The prison system is outdated. No civilized society should punish its criminals: it should rehabilitate them.
- (2) Most university degrees are theoretical and do not prepare students to the real world. They are therefore of very little use.
- (3) Crime does not pay.
- (4) There is no place for censorship in the modern world.
- (5) A man/woman's financial reward should be commensurate with their contribution to the society they live in.

3.2 Identification of errors/infelicities in the corpus using the WordSmith Tools

Wordsmith Tools (Scott 1996), is an integrated package of text analysis programs designed to examine how words behave in texts. The package includes six tools, each for a specific text analysis task. In the present study, two major tools were used. The first was *Wordlist*, which generates word lists in alphabetical and frequency order, as well as statistics such as the total number of words, number of different words, type/token ratio, length of words, number of sentences and length of sentences, so that one can compare texts lexically

¹¹ The 24 essays collected at UFSC, with the help of Professor José Luiz Meurer, were handed in as manuscripts and later transformed into electronic format. Tony Berber Sardinha PUC/SP kindly sent through email the other 43 essays from PUC/SP, UFU, and UNICSUL.

(see Appendixes A and B). The second was *Concord*, which searches a collection of texts stored on computer and displays all the instances of a chosen word or words in their contexts (see Appendix C for a sample concordance of *at*).

Wordlist was used as a starting point to determine which prepositions were going to be the focus of the present study. Among the 100 most frequent words, the ten¹² most frequent prepositions were chosen, as displayed in Table 2:

Rank	Preposition	Tokens
2	TO ¹³	1192
4	OF	935
7	IN	717
16	FOR	268
27	WITH	167
28	ON	164
38	BY	119
40	ABOUT	112
56	AT	84
57	FROM	84

Table 2 - The 10 most frequent prepositions in the Br-ICLE

Table 2 indicates that prepositions are a very frequent morpho-syntactic category in the learner corpus, ten of them appearing among the sixty most frequent words. Although the word *to* had 1.192 occurrences, its usage as a preposition accounts for only 280 occurrences (which would yield a ranking below that of the preposition *in*). Concordancing lines were used to investigate every occurrence of the word *to*, and the whole corpus was POS-tagged¹⁴

¹² Only the 10 most frequent prepositions were chosen because of the size of the corpus. Altogether, these words amount to 3842 instances, which had to be carefully analyzed over and over one by one. This in itself was already a big task, and anything beyond this scope would have been too pretentious for an MA thesis.

¹³ The word *to* was used as a preposition in only 280 instances.

¹⁴ POS tagging (part-of-speech) is a process of attaching a word category tag - often complemented with a series of attributes - to each word in a text. For a thorough discussion of POS tagging on learner corpora see Meunier (1998). Other on-line POS tagging facilities can be found at Lancaster University (CLAWS tagger) Wilson@lancaster.ac.uk, Brill tagger by anonymous ftp blaze.ca.jhu.edu/pub/brill/Programs, ENGCG POS tagger and parser at engcg-info@ling.helsinki.fi, TOSCA POS tagger at tosca@let.kun.nl, and the Xerox POS tagger at <http://www.xerox.com/lexdemo/xlt-overview.html>.

with the TOSCA-tagger¹⁵ (see Appendix D for print screen shot of concordance lines of to_PREP).

By making use of POS-tagged corpora it was possible to determine the total number of prepositions in the corpus, as well as the morpho-syntactic categories which most frequently collocate with these prepositions. A fully POS-tagged corpus can be searched using text retrieval systems such as WordSmith Tools, which provides concordance lines for individual tags in a matter of seconds. Within the *Concord* tool (Scott, 1996), the user can find the 're-sort' command. The point of re-sorting a text is to find characteristic lexical patterns. It can be hard to see overall trends in concordance lines, especially if there are lots of them. By sorting them we can separate multiple search words and examine the immediate context to left and right. For example, we may find that most of the entries may have "at the" or "at a" or "at my" just before the search word, which becomes clearer sorting by the first and second word to the right of the search word.

The next stage consisted of analyzing the ten prepositions through concordance lines in order to separate likely errors/infelicities. These concordance lines were saved in separate files and then submitted to two native speakers of English (one American and one British), both university-level English teachers. The reason for choosing native speakers of two different English varieties is the differences in use prepositions may have. Quirk et al. (1985, p.677) bring to our attention that a word like *school* can be found following these three constructions:

[1] *Sid is at school.*

[2] *Sid is in school.*

[3] *Sid is in the school.*

¹⁵ The TOSCA-tagger allows various tagging schemes. For the sake of simplicity, the corpus was tagged in word_tag format.

He states that the meaning 'enrolled in' is expressed by [1] in British English and by [2] in American English; the meaning 'at the place, not at home' is expressed by [1] and [2] in British English, and by [1] in American English; and the meaning 'within the building' is expressed by [3] in both British and American English.

Whenever an instance was found acceptable to either NS, it was not counted as an error. After the acceptability judgments given by the two NS, a corpus of 283 errors was culled and made ready for categorization. The Université Catholique de Louvain Error Editor (henceforth UCLEE) was used to categorize the errors.

3.3 Error Categorization

The UCLEE, purchased from the University of Louvain, comes with an Error Tagging Manual (Dagneaux et al., 1996) which explains how and why a specific tag is attached to a specific type of error. Based on a corpus of non-native essays, different categories of errors have been defined and a code has been assigned to each one of them. The first letter of the tag indicates the error category: X stands for leXico-grammar, L for Lexis, W for Word order, R for Register, F for Form, and G for Grammar. The manual also provides some principles which the researcher should adhere to when tagging errors (pp.5-7):

Principle 1: Place the tag immediately before the error (word or phrase) that needs to be corrected.

e.g. *At Florianópolis you may find many beaches.*

Correction: *In Florianópolis you may find many beaches.*

The tag (LS) (misuse of independent preposition) has been inserted immediately before *At*:

(LS) At \$In\$ Florianópolis you may find many beaches.

Principle 2: Place the correction immediately after the erroneous word/phrase. For retrieval purposes, the corrected form is preceded and followed by a \$ sign. If there is more than one possible correction, choose the most plausible one.

e.g. *(LS) in \$at\$ home.*

Principle 3: When there are two types of error in the same word/phrase, double tag this word/phrase.

e.g. *she is not very (XADJPR) (XPRCO) good for swim \$good at swimming\$*

Principle 4: Use the 'zero' (0) after the tag to indicate a missing preposition and then supply the missing word between the sign \$...\$. Use the 'zero' (0) between the \$..\$ after the overused preposition to indicate the overuse of this specific preposition.

e.g. *She reads a magazine or a newspaper at least once (WR) in \$0\$ a week. (overuse)*

Thousands (WM) 0 \$of\$ years ago (underuse)

The system developed at Louvain involves a number of steps. First, the learner data is corrected manually by the native speaker(s) of English. Next, the analyst assigns¹⁶ to each error an appropriate error tag and inserts the tag in the text file with the correct version. The inserted correct form should be therefore viewed as 'one possible correct form' - ideally the most suitable one - rather than the one and only possible form.

When the process is finished, the error tagged files can be analyzed using retrieval software tools (e.g., *Concord*), thereby making it possible to count errors, retrieve lists of specific error types, view errors in concordance lines, etc (see Appendix E for print screen shot of concordance lines bearing the error tag XVPR using *Concord*).

Dagneaux et al. (1998) make the point that the purpose of the UCLEE is "to ensure consistency of analysis" (p.166). It should make it possible for different researchers to produce analogous analyses. A categorization in terms of source of errors (e.g. LI transfer), they claim, was rejected because of the high degree of subjectivity involved.

The different error categories and sub-categories employed in the current study will be thoroughly discussed in section 3.3. Examples extracted from the learner corpus will be displayed to exemplify each error category.

3.3.1 Lexical errors (L)

This general category deals with errors involving the semantic (conceptual or collocational) properties of prepositions. In the present study this category is divided into two sub-categories: Lexical Single (LS) and Lexical Phrase (LP).

A) Lexical Single (LS)

(LS) refers to incorrect uses of independent prepositions only:

(LS) in \$at\$ the beginning of next month

B) Lexical Phrase (LP)

(LP) refers to multi-word prepositions, fixed expressions, and possessives involving the preposition *of*:

so if you are (LP) out from \$out of\$ university

they do not support the government (LP) in any means \$by any means\$

(LP) the relationship of a couple \$a couple's relationship\$

3.3.2 Dependent Prepositions (X...PR)

This category includes all errors involving dependent prepositions. The largest groups involve incorrect prepositions with nouns and verbs. As mentioned in Principle 1, the tag is placed in front of the problematic word. The sub-categories are adjective used with the wrong dependent preposition (XADJPR); noun used with the wrong dependent preposition (XNPR); and verb used with the wrong dependent preposition (XVPR).

A) Adjective used with the wrong dependent preposition (XADJPR)

The following are examples of (XADJPR) errors culled from the present corpus.

the teacher was not so (XADJPR) interested on \$interested in\$ teaching

¹⁶ See Appendix F for print screen shot of the error tagging process using the UCLEE.

B) Nouns used with the wrong dependent preposition (XNPR)

The following are examples of (XNPR) errors culled from the present corpus:

Is there any (XNPR) purpose on \$purpose in\$ it

C) Verbs used with the wrong dependent preposition (XVPR)

The following are examples of (XVPR) errors culled from the present corpus:

if you do not (XVPR) think in \$think about\$ any of these subjects

3.3.3 Complementation errors (X...CO)

This category is for complementation errors involving prepositions. For retrieval purposes, the tag is also inserted in front of the complementized word and not in front of the erroneous word. The sub-categories are erroneous complementation of nouns (XNCO); erroneous complementation of prepositions (XPRCO); and erroneous complementation of verbs (XVCO).

A) Erroneous complementation of nouns involving a preposition (XNCO)

Here are some examples of (XNCO) errors collected in the present corpus:

he decided to wait for the right (XNCO) time of doing \$to do\$ it

B) Erroneous complementation of prepositions (XPRCO)

The following are examples of (XPRCO) errors culled from the present corpus:

there are people who are interested (XPRCO) in get \$in getting\$ more money,

C) Erroneous complementation of verbs involving prepositions (XVCO)

Here are some examples of (XVCO) errors collected from the present corpus:

We could (XVCO) ask to ourselves \$ask ourselves\$ why

The UCLE Editor attributes the tag (XVPR) - wrong dependent preposition - for examples like the ones above. Among the examples Dagneaux et al. (1998) provide in the manual we can find '*each European should (XVPR) ask to \$ask\$ himself*' (p. 19). In the

present study I have attributed the XVCO tag to this type of error for the fact that it is not a *wrong* preposition that is being attached to the verbs, but rather, a *wrong* complementation of the verbs. Most verbs included in this category are usually complemented by an object rather than a preposition.

3.3.4 Word Redundant (WR) - overuse errors

(WR) involves all unnecessary uses of prepositions:

most (WR) of \$0\$ people get very depressed in the winter

The examples above show that an unnecessary preposition has been attached to the sentence. The tag \$0\$ indicates that the preposition that precedes the tag is being overused. In case there is more than one word being overused (e.g. *Most (WR) of the \$00\$ people enjoy working out*) the tag \$00\$ is assigned after the words to indicate that both of them are being overused.

3.3.5 Word Missing (WM) - underuse error

This category is for errors involving the omission of a preposition:

They are not paying attention (WM) 0 \$to\$ what the person on TV is saying

A '0' is placed after the (WM) tag to indicate that a preposition is missing. Following the '0' tag, the researcher should supply the missing preposition and place it inside the sign \$...\$.

3.4 Checking doubtful cases through the British National Corpus (BNC)

Notwithstanding native speaker intuitions, errors are not easily recognizable. There are, in fact, great problems in defining error, and considerable variation is found even among native speakers in error identification. For instance, Hughes and Lascaratou (1982) presented thirty-two erroneous and four correct sentences to a panel of thirty judges, ten of whom were Greek teachers of English, ten native speakers of English, and ten native-speaker non-teachers. They found that one of the correct sentences (*Neither of us feels quite happy*) was

judged erroneous by two Greek teachers, and five of the non-teacher native speakers. Another of the correct sentences, which was taken from the *Oxford Advanced Dictionary of Current English*, namely, *The boy went off in a faint*, was judged erroneous by two Greek teachers, nine native speaker teachers, and nine nonteacher native speakers.

In light of this, the instances judged by the two native speakers as neither fully nativelike nor fully erroneous, were checked using the 100 million-word BNC¹⁷ corpus.

The BNC corpus, a collection of British written and spoken language, is one of the areas of research of the University Centre for Computer Corpus Research on Language (UCREL), a research center of Lancaster University directed by Roger Garside. The BNC offers an on-line service which can be used from anywhere in the world. This service allows anyone with access to internet to search for words, phrases, or patterns in the BNC via a simple web interface. The restricted search interface will not return more than 50 hits for each inquiry, with a maximum of one sentence of context for each.

The examples regarded as not fully erroneous, but unlikely to be produced by the native speaker counterparts, constituted the main focus of this stage of the research. Examples involving the possessive use of the preposition *of*, namely, *routine of life*, *standard of life*, and *institution of marriage*, were found in the BNC with respective frequencies of nine, seventeen and twenty-one occurrences out of 100 million words. These frequencies can be contrasted to what the American judge signaled would be the most common way of expressing these ideas in American English, namely, *daily routine*, *standard of living*, and *marriage institution*. The output of the search displayed zero occurrences for *marriage institution*, 361 occurrences for *standard of living*, and 91 occurrences for *daily routine*, which seemed to corroborate her intuitions except for *marriage institution*, which may be more frequent in American English than in British English.

¹⁷ <http://www.comp.lancs.ac.uk/computing/research/ucrel/>

Another doubtful case regarded the cluster *attracted by*, which appeared in the learner corpus in the following sentence: "*In a research made in the United States, they found out that adult people are also attracted by violence.*" Both the present investigator and the American and British judges regarded this as an error, agreeing that the preposition which was to follow the participle verb *attracted* should be the preposition *to*. However, a search in the BNC corpus showed 291 instances of *attracted by* compared to 535 occurrences of *attracted to*. It seems quite difficult to predict, in a sentence like the one above, if '*they*' are *attracted to violence*, or if '*violence*' *attracts them*, which on balance appears to be the same thing. This particular case was not regarded as an error.

The few examples above show how difficult error identification can be, but they also show that nowadays there are many assets the researcher can make use of when trying to acquire a better understanding of the language he/she is analyzing. The BNC has proven to be one of them. Having access to language in its context seems to be one of the most useful ways to achieve such understanding.

3.5 Statistical Procedures

Two statistical tests were put to use in the present corpus in order to ensure more reliable results and also as a means to provide initial indications of significance of results and possible correlations between some of the variables under investigation.

The frequencies of ten prepositions collected from the Br-ICLE were compared to similar corpora derived from the ICLE. All corpora shared several common features: they were written by undergraduate university students of English either in their third or fourth year of university; no student contributed more than 1.000 words; and all of them wrote their essays on argumentative topics.

The X^2 (CHI-SQUARE)¹⁸ statistical test was used to show prepositions whose frequency distribution across the Louvain Corpus of Native English Essays (LOCNESS)¹⁹ and the Brazilian Portuguese subcorpora is statistically significant. Values of X^2 are known to be unreliable for items with expected frequency lower than 5 (see Dunning 1993, p. 5), and possibly result in overestimates for high frequency words and when comparing a relatively small corpus to a much larger one. For this reason, I have randomly extracted 33.750 words from the LOCNESS, and have also used the log-likelihood (G^2) value (Dunning 1993) that does not suffer the same problems as X^2 does with unbalanced sample sizes and high frequency words.

3.6 Research Questions

The research questions addressed in the present study are the following:

1. Is error concentrated on a particular preposition?
2. Is error concentrated on particular error categories?
3. Which semantic/syntactic use is the most problematic for each one of the ten prepositions under investigation?
4. Are the spatial and temporal uses of *in*, *on*, and *at* still problematic for the subjects even though they are assumed to have acquired these structures earlier on in their learning process?
5. Do the advanced learners from UFSC differ significantly as to the distribution of their errors over various categories compared to the advanced learners of PUC/SP, UFU, and UNICSUL?
6. Are the most frequent prepositions - *of*, *in*, *to* - the most frequently misused?

¹⁸ For a thorough discussion of chi-square tests see Rayson & Garside (1998), Kilgarriff (1997; 2001), and Dunning (1993).

¹⁹ ICLE contains a collection of texts written by British and American students and which can be considered a referential material for comparisons of underuse or overuse of features in learner essays.

7. Contrasted to EFL corpora of different language backgrounds, does the present corpus display similar frequencies of prepositions? Which prepositions are significantly under- and overused compared to the LOCNESS?

3.7 Conclusions

According to Abbot (1980, p. 122) "one major requirement of any process rigorous enough to be called an analysis is that the results should be verifiable by other scholars using the same procedures". This is exactly the underlying principle in the present chapter. It should make it possible for different researchers to produce analogous analyses. A word of explanation is necessary as to why only one syntactic feature (i.e., prepositions) was selected for scrutiny. Many early EA studies have dealt with an assortment of grammatical errors, which have led to fuzzy error categorizations and method descriptions. In the present study, however, the assumption was that if the syntactic variety of the data were restricted to only one grammatical unit, it would be easier to focus, at each stage of the analysis, upon the procedural matters that were viewed as one of the main concerns of the study.

In the next chapter I shall discuss correct and incorrect syntactic and semantic uses of each preposition. By doing this, I hope to indicate major patterns of correct and incorrect uses and their respective distribution of frequency in the corpus, as well as to propose an explanation for the different error categories in terms of 'source'.

CHAPTER FOUR

RESULTS AND DISCUSSION

4. Introduction

This chapter will be divided in two main parts: (a) sections 4.1 to 4.10 describe all the major patterns of correct and incorrect use for each preposition; (b) section 4.11 addresses the research questions outlined in chapter 3 taking into account what has been described from sections 4.1 to 4.10.

Regarding the correct uses of these prepositions, a description in terms of syntactic and semantic use is put forward in relation to the most common contexts where they have occurred (I do not cover the entirety of correct uses of each preposition). As for the incorrect uses, all of the errors and their respective error categories are described for each preposition and *suggestions* are given for their likely sources and causes. For practical reasons, contrary to normal practice, that is, to display the errors at the end of the study, usually in an Appendix, it was judged better to instantiate the errors along with their description so as to facilitate reading.

Section 4.11 addresses the research questions put forward at the end of chapter 3. Most of the answers to these questions have emerged in the analysis of sections 4.1 through 4.10.

4.1 The Preposition ABOUT

The preposition *about* occurred 112 times in the Br-ICLE, 99 occurrences (88%) being regarded as correct and 13 as incorrect. Before describing the most frequent errors, their categories, and their likely sources, I shall point out some of the uses with which the subjects did not have problems.

4.1.1 Correct uses of ABOUT

Table 3 displays the distribution of correct uses of *about* in relation to verbs, nouns, adjectives, interrogatives (i.e., questions words such as *how* and *what*), and phrasal verbs.

Semantic and syntactic uses	Number of correct uses	%
Verbs	56	57
Nouns	16	16
Adjectives	15	15
Interrogative	5	5
Phrasal Verbs	2	2
Others	5	5
TOTAL	99	100%

Table 3 – *About*: distribution of correct uses

Table 3 shows an overwhelming use of *about* in adjunct position (i.e., following verbs), which corroborates what most grammars imply about its syntactic use. Here are some verbs followed by *about* (frequencies between parenthesis):

care (7) do (1) know (3) worry (3)
 decide (1) talk (7) decide (1) teach (2)
 dream (1) think (19) hear (2) forget (1)

In most of the verbs above, the meaning the subjects wanted to convey was that of *concerning* something or someone. In fact, by looking at the nouns which preceded the preposition *about*, the meaning *concerning* is the most commonly used as well. Here are some nouns followed by *about*:

article (1) information (1) option (1) story (1)
 essay (2) expectation (1) position (1) report (1)
 message (1) knowledge (2) problem (1) misunderstanding (1)

Most of these nouns convey the notion that something is going to be characterized after them, that is, the object of concern is going to be revealed in the prepositional phrase. Swan (1980, p. 16) points out that *about* instead of *on* is used "when the information given is more general, or the style of communication is more casual".

Regarding the noun *knowledge*, which was the object of many discussions between the present investigator and the two native speakers, the *Longman Dictionary of Contemporary English* attributes the preposition *of* as the preposition to follow the noun *knowledge*. Nonetheless, on its next entry the adjective *knowledgeable* is described as having *about* as the dependent preposition. So, we have *knowledge of* something, but we are *knowledgeable about* something else. Searching the BNC for *knowledge about* and *knowledge of* corroborated the native speakers' intuitions, who claimed that *knowledge about* was also 'acceptable'. A total of 385 instances were found for *knowledge about*, a very small number compared to the 4151 occurrences of *knowledge of*. Despite the huge difference, both native speakers agreed that the contexts where *knowledge about* appeared (e.g., ...*knowledge about grammar*..., ... *knowledge about spiritual things*..., *knowledge about the real world*) were acceptable according to their standards.

Following the same semantic use employed with verbs and nouns, adjectives followed by *about* were also commonly used with the meaning *concerning*. Here is a list of some adjectives followed by *about*:

worried (8) excited (1) careful (1) right (1) concerned (3) secure (1)

Finally, there were three other semantic uses correctly employed by the subjects. The first, regards the use of interrogative pronouns such as *what* and *how* followed by the preposition *about* to inquire or suggest something. The second refers to the phrasal verb *bring about* conveying the idea of 'make happen', as in *Therefore, the death penalty would bring about a lot of problems*; and the third, to indicate degree: *men watch television about forty hours a week*.

The next section deals with the assortment of errors involving the preposition *about*, how they were spread within the error categories, and their likely sources.

4.1.2 Incorrect uses of ABOUT

Table 4 provides the frequency of errors within different error categories involving the preposition *about*:

Error categories	Number of incorrect uses	%
XNPR	5	3
XADJPR	3	2
WR	2	1
XVPR	1	8
XPRCO	1	8
LS	1	8
TOTAL	13	100%

Table 4 – *About*: distribution of errors

Table 4 shows that the preposition *about* was wrongly employed as a dependent preposition on nouns, adjectives, and verbs (XNPR, XADJPR, XVPR) in most of the errors (9 occurrences). When dependent on nouns (XNPR), *about* accounted for 38 per cent of the errors. Here are the cases where *about* was wrongly employed with nouns (the suggested correction is placed to the right of the examples):

different points of view and <u>approaches about</u> political subjects ...	[approaches to]
there has been some <u>criticism about</u> the relation between ...	[criticism of]
Through a <u>research about</u> cells ...	[research on]
Recently a <u>research about</u> television was done ...	[research on]
such a perfect <u>satire about</u> the Stalinism ...	[satire of]

The errors above seem to suggest that the subjects have generalized *about* for the Portuguese preposition *sobre*. As it can be noticed, the semantic role attributed by *about* indicates *concern*, which in turn, can also be attributed by the Portuguese preposition *sobre*. Thus, it might be inferred that the errors above might stem from L1 transfer, in that the subjects seem to have attributed the preposition *about* as an equivalent of the Portuguese preposition *sobre* (meaning *concerning*).

About was misused as dependent on adjectives (XADJPR) in 23 per cent of the errors. However, with adjectives, the semantic meaning *concerning* did not seem to derive from the Portuguese preposition *sobre*. It is the Portuguese preposition *de* that the subjects seem to be

generalizing as an equivalent of *about*, except for the second sentence where *com* seems to be what the subject had in mind.

to be more <u>conscious about</u> himself, <u>about</u> his possibilities ...	[conscious of]
she gets very <u>impressed about</u> the play "Macbeth" ...	[impressed with]
they are usually living <u>scared about</u> everyone ...	[scared of]

The other errors were spread among four error categories, namely, word redundant (WR), preposition dependent on verbs (XVPR), lexical single (LS), and wrong complementation of preposition (XPRCO) as the examples below indicate (the error category tag and the suggested correction are placed to the right of the example):

What <u>about start</u> this modification ...	[XPRCO - about starting]
several people are questioning <u>about</u> the ethics of today's press ...	[WR - 0]
magazines are always reporting <u>about</u> politics ...	[WR - 0 / LS - on]
to <u>reflect about</u> the reading ...	[XVPR - reflect on]
the wide domain <u>about</u> modern multimedia techniques ...	[LS - of]

Overall, the subjects correctly used the preposition *about* 88 per cent of the time, which reveals that it does not seem to be a troublesome preposition for them. Both errors and correct uses of *about* seem to indicate that the most prototypical meaning the subjects have for this preposition is the one of *concerning* something or someone. Most of the errors involving *about* seem to stem from L1 transfer, where the subjects seem to have generalized *about* as an equivalent preposition of the Portuguese preposition *sobre*, and at times (specially with adjectives), with the Portuguese preposition *com* and *de*, both meaning *concerning*.

4.2 The preposition AT

At accounted for 84 occurrences in the Br-ICLE, 70 occurrences (83%) being regarded as correct and 14 occurrences as incorrect. Before describing the most frequent errors, their categories, and their likely sources, I shall point out some of the uses with which the subjects did not have problems.

4.2 Correct uses of AT

As expected, many uses of *at* involve adverbial uses, place and time relations, and its dependent use on verbs. Table 5 displays the distribution of correct uses of the preposition *at* among these semantic and syntactic categories:

Semantic and syntactic uses	Number of correct uses	%
Adverbial	25	3
Place	19	2
Time	17	2
Verbs	6	9
Others	3	4
TOTAL	70	100%

Table 5 – *At*: distribution of correct uses

The high frequency of adverbial uses constructed with *at* was rather surprising. The adverb of intensity *at least* alone, accounted for 19 per cent of all occurrences of *at*, followed by *at all* (6 occurrences), and other adverbs such as *at once*, *at present*, *at the best*, and *at last* which occurred once each. Altogether, adverbial uses of *at* accounted for 36 per cent of its correct use, their high frequency seemingly resulting from the genre of the topics the subjects were given to choose from, (i.e., argumentative essay writing), and also showing that these expressions are learned as chunks.

The use of place and time relations expressed by *at* also figures prominently, accounting for 51 per cent of the correct uses of this preposition. Below are some examples of place and time expressions involving *at* with their respective frequencies in parenthesis:

at home (5) at school (5) at university (3)
 at the moment (4) at night (3) at the beginning/end (3)

The verb dependent uses of *at* regarded the verbs *look* (4), *glance* (1), and *repent* (1), which together accounted for 9 per cent of its correct use.

The next section deals with the assortment of errors involving the preposition *at*, how they were spread within the error categories, and their likely sources.

4.2.2 Incorrect uses of AT

Table 6 indicates that the errors concerning *at* were distributed among its independent use (LS), verb dependent (XVPR), overuse (WR), underuse (WM), and noun dependent (XNPR):

Error categories	Number of incorrect uses	%
LS	8	5
XVPR	2	1
WR	2	1
WM	1	7
XNPR	1	7
TOTAL	14	100%

Table 6 – *At*: distribution of errors

Table 6 shows that LS errors figure prominently in relation to the other error categories. These errors involve the use of *at* in place and time expressions as indicated by the examples below:

Here, <u>at</u> Florianópolis ...	[in]
<u>At</u> Argentina, ...	[In]
the terrible life style that we faced <u>at</u> São Paulo ...	[in]
job's future, as we know it today, is <u>at</u> a check mate position ...	[in]
<u>At</u> the 60's ...	[In]
<u>At</u> the end, the sentence is given ...	[In]
would not raise votes <u>at</u> the political elections ...	[in]
The new regime, democratic <u>at</u> the beginning, ...	[in]

Interestingly, but rather expectedly, in all of the errors where *at* was used as an independent preposition, *in* is the target preposition. This very fact appears to point to what was described in Chapter 2, where a contrast was made between the Portuguese and the English prepositional systems, and it was pointed out that the Portuguese preposition *em* can be expressed in English through *at*, *in*, and *on*, a view also reinforced by Swan and Smith (1990, p. 99). Besides, if any of the above examples were translated into Portuguese, the preposition to be used would certainly be *em*.

The other errors were scattered among five error categories which are listed below (the error category tag and the suggested correction are placed to the right of the example):

<u>At</u> this way, the media can transfers ...	[WR - This way]
but <u>at the</u> backstage the situation is ...	[WR - backstage]
Rita <u>enrolls at</u> a program called Open University ...	[XVPR - enrolls in]
colleagues that <u>graduated at</u> colleges ...	[XVPR - graduated from]
participate in all events of society aiming <u>0</u> its development ..	[WM - aiming at]
will demand <u>knowledges at</u> math ...	[XNPR - knowledge of]

Even though the errors above reflect other than place and time misuses, they too, resemble (LS) errors in that translating these sentences into Portuguese, with the exception of *At this way*, and *aiming at its development* the target preposition is *em*.

Summing up, by making use of EA and CA it was possible to map out all of the semantic and syntactic correct and incorrect uses of *at*. It is proposed that these errors may stem from the fact that the Portuguese preposition *em* has as English equivalents the prepositions *in*, *on*, and *at*, which can be used to convey very similar meanings, (i.e., time, space).

4.3 The preposition BY

By accounted for 119 occurrences in the Br-ICLE, 99 occurrences (82%) being regarded as correct and 20 occurrences as incorrect. Before describing the most frequent errors, their categories, and their likely sources, I shall point out some of the uses with which the subjects did not have problems.

4.3.1 Correct uses of BY

The preposition *by* was chiefly used after past participle verbs to indicate the agent of an action in passive voice. Other correct uses of *by* involve time expressions, to indicate the means by which something was done, fixed expressions, and as an explicatory device. Table 7 displays the distribution of these correct uses:

Semantic and syntactic uses	Number of correct uses	%
Past participle verbs	59	6
Time	8	8
Manner	9	9
Fixed expressions	7	7
Explicatory	2	2
Others	14	1
TOTAL	99	100%

Table 7 – By: distribution of correct uses

By was used after past participle verbs to indicate the agent of an action in 59 occurrences (60%) of its correct use, as demonstrated by the examples below:

But not from the same reasons given by Eric Fromm.

The best reason considered by the government is that they ...

Temporal expressions involving *by* accounted for only 8 per cent of its correct use and involved expressions like *as time goes by*, *by the end of* + noun, *by the time* + pronoun, *by next century*, and *years went by* (where *by* was used as a prepositional particle).

By also appeared in a few fixed expressions such as *day by day*, *by law*, *by means of*, and *by the way*. This use of *by* accounted for only 7 per cent of its correct use.

Another correct semantic use employed by the subjects concerned its use to indicate the *means* someone used to achieve an end. In this particular context, which accounted for 9 per cent of its correct use, *by* was usually employed after an adjunct followed by an '-ing' clause, as demonstrated in the examples below:

This fact can be confirmed just by analyzing how many students have passed ...

Television keeps on destroying whole homes by 're-formatting' the time which people should... couldn't he support all Germany alive by taking the Jewish money ...

Other contexts in which *by* was correctly employed were to explicate a point of view, as in *By social/psychological problems I mean ...*, to indicate degree, as in *unemployment rose by 18% last year*, to indicate the creator of something, for example, *a song by Caetano Veloso*, and to indicate that someone did something without help: *they can think by themselves*.

Altogether, these semantic uses accounted for 16 per cent of the correct uses of the preposition *by*.

The next section deals with the assortment of errors involving the preposition *by*, how they were distributed within the error categories, and their likely sources.

4.3.2 Incorrect uses of BY

The errors involving *by* encompass its use as an independent preposition (LS), as dependent on verbs (XVPR), as the wrong complement of a noun (XNCO), its avoidance (WM), and its incorrect use in fixed expressions (LP). Table 8 displays the distribution of errors among the categories listed above:

Error categories	Number of incorrect uses	%
XVPR	8	4
LS	6	3
LP	3	1
XNCO	2	1
WM	1	5
TOTAL	20	100%

Table 8 – *By*: distribution of errors

In the cases where *by* was used as the wrong verb-dependent preposition (40%), most of the verbs were in the past participle and the semantic uses intended by the subjects were varied, as attested in the examples below (the suggested correction is placed to the right of the example):

a black man <u>indicted by</u> double homicide ...	[indicted for]
we should use the ones <u>made by</u> glass or metal ...	[made of]
he is <u>valued by</u> what he has ...	[valued for]
someone who is <u>suffering by</u> cancer ...	[suffering from]
they <u>swear by</u> God ...	[swear to]
while he would to be <u>valued by</u> what he is ...	[valued for]
a person who is <u>suffering by</u> any disease ...	[suffering from]
Brazilians are <u>known by</u> their hospitality ...	[known for]

Contrasting the Portuguese and the English prepositional systems appeared to reveal that *by* was being generalized in most cases for the Portuguese prepositions *por* and *de*, which

in fact, are used in Portuguese, in some cases, to indicate the same semantic meanings the subjects were trying to convey in English, (i.e., cause, agent, content, etc). Notice that the suggested corrections employ four different English prepositions as the ones the subjects should have used: *for*, *to*, *from*, *of*. Seeking their equivalences in Portuguese, we find that all of them can be translated either by *de* or *por* (refer to Chapter 2 pages 53-54 for examples). Therefore, it is my contention that the errors evidenced above are very likely to stem both from multiplicity of forms of both prepositional systems, especially because *by*, *for*, and *to* are equivalents of *por*, and *from* and *of* are equivalent of *de*, and because the preposition *by* is chiefly used with past participle verbs and some adjectives, which in turn, may lead students to generalize its use with these kinds of verbs and adjectives.

Independent uses of the preposition *by* were also largely wrongly employed by the subjects (30%). The errors below seem to indicate that the subjects have not distinguished *by* from *through* and have used it as an equivalent of the Portuguese preposition *através*. On the other hand, the errors in which *in* and *for* are the likely target prepositions, the subjects might have generalized the common use of *by* with both past participle verbs and adjectives as the correct way to express the meaning they wanted. Thus, regarding independent uses of the preposition *by*, the errors seem to stem both from L1 transfer and overgeneralization of the common use of *by* with past participle verbs and adjectives.

tolerance cannot be dealt with <u>by</u> a compelling education ...	[through]
killed <u>by</u> the most cruel manners ...	[in]
In Peru, Fujimori is trying to be reelected <u>by</u> the third time.	[for]
how much does it cost to become a little bit young <u>by</u> plastic modelling.	[through]
but where you will buy the products showed <u>by</u> commercials ...	[in]
Since plastic surgeries became popular <u>by</u> the TV and the newspapers ...	[through]

With respect to fixed expressions, the preposition *by* was misused in the following contexts: *she decided by her own*; *we must begin by the bottom to reach the top*; and *that happened by random*. As to *by her own*, it is very likely that this error stems from the fact that

we use *by herself* meaning 'without help'. As regards *by random*, further context is required to point out the likely source of such error: *When all citizens are able to share the benefits of such things or facts, it is merely something that happened by random while the real important promises regarding equality and social justice are forgotten or set apart for future discussions*. It appears that what this subject had in mind was the idea of *by chance* which is similar to the meaning conveyed by *at random* (= without any plan). So, the problem here might derive from the use of *by* to indicate that the event was not planned, disregarding or not knowing its collocation (*by* does not collocate with *random*). Quite likely, what the subject wanted to say was either *by chance*, where we would have a lexical rather than a lexicogrammatical error, or *at random*, where *random* is correct but the preposition preceding it is not. Alternatively, the subject could have used the adverb *randomly* instead of *by accident* or *at random*. This reinforces the fact that EA can be blurred at the edges, and that attributing causes to errors is a very subjective task. Therefore, the reader should view the sources of the errors suggested here, as *probable* sources, since the *real* source may never be encountered.

The other errors involved the wrong complementation of the noun *wish* as in *the wish by learning*, which was produced twice by the same subject, and the omission of *by* (represented by "0") in the sentence *they motivate them to buy women's clothing and shoes "0" saying they look nice wearing that sort of clothes*, where the subject failed to provide the preposition *by* to indicate the means used to motivate the purchase of women's clothes and shoes.

The preposition *by* was misused nearly one fourth of the time it was employed. The errors were distributed mostly among verb-dependent uses of *by* and its independent use as a preposition which together accounted for 70% of the total number of errors. It is argued that in most cases the likely sources for these errors were either problems of equivalence between the prepositional systems of the two languages, or overgeneralizations regarding the use of *by* with past participle verbs and some adjectives.

4.4 The preposition FOR

For accounted for 268 occurrences in the Br-ICLE, 251 occurrences (94%) being regarded as correct and 17 occurrence as incorrect. Before describing the most frequent errors, their categories, and their likely sources, I shall point out some of the uses with which the subjects did not have problems.

4.4.1 Correct uses of FOR

Among the correct uses of the preposition *for*, eight syntactic and semantic uses figure prominently: *for* was used as a logical connector, as dependent on adjectives, nouns, and verbs, *for* was also used to indicate temporal relations, as a prepositional particle, to indicate who benefited from something, and to indicate the reason for an action.

Semantic and syntactic uses	Number of correct uses	%
Logical connector	50	2
Adjectives	45	1
Nouns	39	1
Verbs	37	1
Time	25	1
Prepositional particle	21	7
Benefaction	13	5
Reason	9	4
Others	12	5
TOTAL	251	100%

Table 9 – *For*: distribution of correct uses

Regarding logical connectors, *for* was used as an additive connector of exemplification in 59 per cent of the uses of logical connectors with *for example* and *for instance* occurring 16 and 13 times respectively. The connector *as for* was employed three times to introduce a topic, and the causal connectors *for this reason* and *for* were used altogether 14 times to indicate either effect, reason, or purpose. The use of logical connectors accounted for 20% of the total number of correct uses of the preposition *for*.

For was used as dependent on adjectives, nouns and verbs in 48% of its correct use.

Here are some adjectives, nouns, and verbs which occurred to the left of *for* (the frequencies are given in parenthesis):

Adjectives

better	(3)	difficult	(2)	distasteful	(1)	eager	(1)	enough	(3)
equal	(1)	essential	(1)	famous	(1)	fundamental	(1)	good	(5)
hard	(1)	important	(6)	necessary	(3)	obvious	(1)	prepared	(1)
responsible	(8)	shameful	(1)	true	(2)	wonderful	(2)		

Nouns

affection	(1)	attention	(1)	challenge	(1)	change	(1)	consideration	(1)
formula	(1)	instrument	(2)	opportunity	(2)	passion	(1)	place	(3)
program	(1)	reason	(4)	respect	(1)	solution	(5)	time	(2)

Verbs

ask	(2)	atone	(1)	claim	(2)	die	(2)	fight	(7)	increase	(1)
judge	(2)	opt	(1)	pay	(3)	prepare	(3)	search	(2)	struggle	(1)
wait	(3)										

Temporal uses of *for* accounted for 10% of its correct use. Here are the most typical examples of how *for* was used syntactically to convey temporality:

for + (the) many/more/some/next/following/first + years/hours/minutes/time

for all of + adjective pronoun + life/lives

For was correctly used as a prepositional particle attached to the verb *look* meaning to seek, or to search, in seven per cent of its correct use. Interestingly, the things that were mostly *looked for* were: jobs, ways of looking younger, and opportunities.

Concerning its semantic use as a way to indicate who or what benefited from an action, *for* was employed in five per cent of its correct usage. Here are some examples extracted from the Br-ICLE which bear witness to this particular semantic use:

a house in which everyone has a TV for him or herself.

social justice for all members of society.

Lastly, *for* was correctly employed four per cent of the time to indicate the reason for an action:

theory was drastically changed for personal interests.
people kill for less than a hundred real.

The next section deals with the assortment of errors involving the preposition *for*, how they were spread within the error categories, and their likely sources.

4.4.2 Incorrect uses of FOR

The errors involving *for* comprise dependent use on adjectives, nouns, and verbs (DANV)²¹, misuses in the complementation of adjectives, nouns, and verbs (CANV)²², independent uses of *for* (LS), and overuse (WR). Table 10 displays the error distribution among the error categories mentioned above.

Error categories	Number of incorrect uses	%
DANP	8	4
CANP	5	2
LS	3	1
WR	2	1
TOTAL	18	100%

Table 10 – For: distribution of errors

The preposition *for* was erroneously used as dependent on adjectives, nouns and verbs (DANP) in 44% of the total errors. Here are the errors:

seeing how big is it and how <u>different for</u> culture context.	[XADJPR - different from] they
say are more <u>relevant for</u> the good of all citizens.	[XADJPR - relevant to]
A good <u>example for</u> this is when we see on the TV ...	[XNPR - example of]
watching television, <u>attracted for</u> a new word .	[XVPR - attracted to, by]
s/he uses television just to <u>look for</u> .	[XVPR - look at]
programs are <u>oriented for</u> people with lower education ...	[XVPR - oriented to]
they are <u>restricted for</u> those who can pay for a cable TV ...	[XVPR – restricted to]

The errors above indicate that the target prepositions intended by the subjects were either *from*, *to*, *of*, or *at*. In the second, fourth, and sixth examples the target preposition the

²¹ For the sake of simplicity, I have lumped together the error categories XADJPR, XNPR, and XVPR under the tag DANV, and the error categories XADJCO, XNCO, and XVCO under the tag CANV.

subjects failed to use *to*. Notice that *for*, *at* and *to* can be translated to *para* in Portuguese. In the first and third examples, the prepositions *from* and *of* were the target. The first example is grammatically poor, being difficult to infer about its causes. In the third example, the Portuguese translations *Um bom exemplo disso/para isso é quando vemos na TV ...* which might explain why the subject opted for using *for* instead of *of*

For was also misused as the complement of adjectives, nouns, and verb. These errors accounted for 28% of the errors involving *for*.

<u>important for preparing</u> yourself to ²³ the....	[XADJCO - important to prepare]
we need <u>money for staying</u> alive.	[XNCO - money to stay]
a good <u>way for improving</u> human relations...	[XNCO - way to improve]
we must <u>give for them</u> a second chance ...	[XVCO - give them]
<u>showing for people</u> that they will never...	[XVCO - showing people]

The first three errors differ from the last two in that the subjects who made the first errors used the structure *adjective/noun + for + v-ing* to attribute the reason for something, whereas the subjects who made the last errors used the structure *verb + for + indirect object* to express intended target, or goal. Even though the errors above differ in relation to their syntactic construction, they resemble in one aspect: they have the Portuguese preposition *para* as an equivalent of *for* being used to express both reason and intended target: *dinheiro para sobreviver* (reason); *mostrando para as pessoas* (intended target).

Incorrect independent uses of *for* (LS), accounted for 17% of the total number of errors involving this preposition. Once again, the Portuguese prepositions *para* and *por* appear to have been generalized as equivalents of the preposition *for*.

the number of plastic surgeries already used <u>for</u> adults and teenagers also.	[by]
It is a belief <u>for</u> several religions that fertility is a gift from God.	[in]
Here I am going to talk about love <u>for</u> a couple's reality.	[in]

Overuse errors (WR) accounted for 11% of the total number of errors.

²² See footnote 21.

²³ This particular error is dealt with in section 4.9.2 - Incorrect uses of TO

For a long time ago, people were walking on the street with their family ... ["0"]
 able to ask for or question why they don't have the same opportunities ... ["0"]

Only six per cent of the occurrences of the preposition *for* were regarded as erroneous, which indicates that the subjects seem to have acquired its use as a logical connector, as dependent on certain adjectives, verbs, and nouns, as a way to denote time, and as a very common particle attached to the verb *look*. The errors were mostly concentrated in its use as a dependent preposition, as well as in the complementation of certain verbs and nouns. Nearly half of the errors involving *for* (8) might result from the fact that the subjects seem to be regarding it as the best equivalent to the Portuguese prepositions *para* and *por*, disregarding, or even not acknowledging, that *para* and *por*, may as well have as English counterparts the prepositions *by*, *to*, and *at*.

4.5 The preposition FROM

The preposition *from* occurred 84 times in the Br-ICLE. It was correctly employed 95% of the time. The following section is a brief account of its correct use. The errors and their likely sources are addressed in section 4.5.2.

4.5.1 Correct uses of FROM

Two semantic uses account for most instances regarding the preposition *from*: source and origin. Here are some examples of these uses:

Source and Origin

<i>fertility is a gift from God ...</i>	[source]
<i>our idea of a pretty face varies from culture to culture ...</i>	[source]
<i>people from the northeastern part of Brazil ...</i>	[origin]
<i>every person which is from Bahia, Ceará, ...</i>	[origin]

From was also used in two instances to indicate time e.g. *from that moment on, a century from now*, and in one instance to indicate range: *plastic modelling increased from 8% to 90%*. The complex preposition *away from* accounted for three occurrences.

4.5.2 Incorrect uses of FROM

The errors involving *from* were distributed within the error categories (LS), (XNPR), and (LP) as shown in the examples below (the error tag and suggested correction are placed between brackets to the right of the examples):

the <u>difference from</u> friends and enemies ...	[XNPR – difference between]
a recent report published in "Revista Já", <u>from</u> "Diário Popular" journal	[LS – of]
Her book called Sex, <u>from</u> 1992, was criticized by the Pope!	[LS – of]
so if you are <u>out from</u> the university environment ...	[LP – out of]

The few errors regarding *from* seem to indicate that the subjects may have confused the use of the prepositions *from* and *of*, since both of them can mean *de* in Portuguese. In the first sentence, the subject wanted to mean either *diferença entre* or *diferença de*. In either case, the noun *difference* only collocates in this type of context with the preposition *between*, whereas in Portuguese, both *de* and *entre* may be used with the noun *diferença*.

The very low percentage of errors involving *from* (5%) is an indicator that for the subjects of this study this preposition does not cause many problems. *From* was employed in the majority of the cases to indicate source or origin, and the few errors involving its use seem to be related to the two ways the Portuguese preposition *de* may be expressed in English, (i.e., *from* and *of*).

4.6 The preposition IN

In accounted for 717 occurrences in the Br-ICLE, 658 being regarded as correct (92%), and 59 as incorrect. Before describing the errors, how they were distributed within the error categories, and their likely sources, the following section will address the instances with which the subjects did not have problems.

4.6.1 Correct uses of IN

Regarding its semantic and syntactic use, *in* was employed to indicate place, time, manner, subject matter, as a part of the structure of many connectors, followed by gerund verbs, and to form the complex preposition *in front of*.

Semantic and syntactic uses	Number of correct uses	%
Place	291	4
Time	70	
Manner	55	
Subject matter	36	
Part of connectors	113	1
Gerund verbs	18	
<i>In front of</i>	10	
Others	65	1
TOTAL	658	100%

Table 11 – *In*: distribution of correct uses

Nearly half of the time the preposition *in* was employed correctly, it was employed to indicate place. The places indicated by *in* varied from very concrete landmarks to more abstract and dimensionless spaces or notions. By looking at the examples below we are able to construct a spectrum which varies from concrete to more abstract places (all the examples were extracted from the Br-ICLE):

people can smoke grass in Amsterdam...

discover new things in a world that ...

what happens in most courses is that ...

things would change in nature too.

in our super competitive job market ...

beauty has an important value in our society ...

youth is in our minds ...

only exists in his dreams ...

CONCRETE



MORE ABSTRACT

Regarding its temporal use, *in* was employed most of the time to describe spans of time, e.g., *in 1999*, *in the last years*, *in the next century*, *in a digital era*. In many instances, *in* was used together with the word *life* in the structures *in + adjective pronoun/genitive + life/lives* or *in life*.

In also accounted for instances which indicated manner: *in love; in haste; in coma; in fear; in absentia; in a nutshell; in a very democratic way; in a selfish and busy way; in miserable conditions of life, etc.*

In was also employed to indicate subject matter, as the examples extracted from the BRICLE attest: *investing in education, graduated in Fashion, a glossary in four languages, Phd in Linguistics, amount of money in gold, etc.*

Many connectors, with various discourse functions, were constructed using the preposition *in*. Below is a list of these connectors with their respective frequencies in parenthesis:

in a way	(2)	in addition	(8)	all in all	(1)	in brief	(1)
in contrast	(1)	in my opinion	(16)	in order to	(56)	in relation to	(2)
in spite of	(4)	in sum	(1)	in terms of	(5)	in view of	(4)
in this way	(2)	in fact	(10)	in other words	(5)		

The other two syntactic uses involving *in*, regarded its use preceding gerund verbs as in

I am very interested in reading books

There is no use in attending an university

There is no sense in teaching a cat

and as a component of the complex preposition *in front of* which occurred ten times in the BRICLE.

4.6.2 Incorrect uses of IN

The preposition *in* displayed an assortment of errors which were divided within ten error categories. Three error categories, namely, (LS), (LP), and (XVPR), accounted for most of the errors, the other errors being scattered within seven other categories as indicated in Table 12:

Error categories	Number of incorrect uses	%
LS	18	3
LP	11	1
XVPR	11	1
Others	19	3
TOTAL	59	100%

Table 12 – In: distribution of errors

The error category (LS) was responsible for 30% of all the errors regarding *in*. Here are the instances where *in* was employed as a wrong independent preposition (the suggested correction is placed inside brackets to the right of the example):

Rita started to be recognized <u>in</u> the university as a good student ...	[at]
prefer to spend their birthdays <u>in</u> a hairdresser's ...	[at]
they are <u>in</u> home ...	[at]
speaking with friends even <u>in</u> the bus stop ...	[at]
<u>in</u> each time people considered love differently ...	[at]
she gets a new job <u>in</u> a bistro ...	[at]
you must live each one <u>in</u> the correct time ...	[at]
he didn't want her <u>in</u> the university ...	[at]
he decides to challenge himself <u>in</u> trying to drown deep ...	[by]
these things showed off <u>in</u> TV are respectable ...	[on]
what would be shown <u>in</u> TV news ...	[on]
<u>In</u> TV things are ready ...	[on]
the jewels paid by my father <u>in</u> the same night ...	[on]
cakes recipes <u>in</u> the first pages of important newspapers ...	[on]
<u>In</u> "Fame", it was possible to notice that ...	[on]
guaranteeing peace <u>in</u> our planet ...	[on]
other examples can be seen <u>in</u> "Fame" ...	[on]
<u>In</u> the teen years, the person is looking for what ...	[During]

Two important aspects can be drawn from the errors above. First, that the target prepositions were most of the time either *at* or *on*. Secondly, by translating most of these sentences into Portuguese we notice that the equivalent for *in* in Portuguese is *em*. It was pointed out in Chapter 2 that at least three prepositions can be used in English as equivalents of *em*, namely, *in*, *on*, and *at*. It was also pointed out that the preposition *em* can also be used to convey both place (e.g., *garantindo paz no nosso planeta*), and time (e.g., *as jóias pagas pelo meu pai na mesma noite*). With the exception of *he decides to challenge himself in trying to drown deep*, which conveys the idea of manner, all the errors were related either to place or

time expressions whose Portuguese equivalent preposition is *em*. It may be argued, then, that these errors may stem from the multiplicity of English prepositions to express place and time relations which, in turn, may be expressed in Portuguese only with the preposition *em*.

The misuse of *in* in lexical phrases (LP errors) accounted for 19% of the total number of errors involving this preposition. Below is the complete list of errors under the error category (LP) (the suggested correction is placed between brackets to the right of the example):

the course <u>in</u> Letras (occurred twice)	[Letras course]
the course <u>in</u> Mathematics ...	[Mathematics course]
the course <u>in</u> Biblioteconomia ...	[Biblioteconomia course]
supported by the government <u>in any means</u> ...	[by any means]
<u>In opposite</u> this situation, ...	[Opposite to]
can put their prestige <u>in risk</u> .	[at risk]
so, <u>in the moment</u> that you find ...	[at the moment]
the person can enjoy their work, and <u>in the same time</u> ...	[at the same time]
these books are often <u>in the top</u> of the list ...	[at the top]
it wasn't what he was searching, <u>in the contrary</u> , ...	[on the contrary]
what is going on <u>in the other side</u> of the world ...	[on the other side]

The first three errors involving the noun *course* were made by the same subject, who seemed to be using *the course in* as an equivalent of *o curso em/de*. Differently from errors involving the independent use of the preposition *in*, these errors reflect its misuse in semi-fixed expressions which are normally learned as word chunks. Notice that in more than half of the errors the target prepositions were again *on* or *at*.

The errors involving *in* as a dependent on verbs, accounted for 19% of the total number of errors for this particular preposition. Interestingly, the verb *think* was the verb with which the subjects had most problems. In Portuguese we can either *pensar em*, or *pensar sobre*, whereas in English the preposition most commonly employed with the verb *think* is the preposition *about*, and less frequently the preposition *of*. Here are the errors concerning *in* as dependent on verbs (the suggested correction is placed between brackets to the right of the example):

if you get <u>concentrated in</u> something ...	[concentrated on]
---	-------------------

if you do not <u>think in</u> any of these subjects ...	[think about]
we can <u>think in</u> "living longer" ...	[think about]
we could <u>think in</u> the religious ...	[think about]
couples believe that is the better option <u>thinking in</u> their behavior ...	[to think about]
Now, <u>thinking in</u> the two points, ...	[thinking about]
this person that <u>thinks in</u> money every time ...	[thinks about]
the free time was <u>transformed in</u> work time ...	[transformed into]
qualified professionals to <u>work in</u> the criminals recovering ...	[work on]
some couples had always <u>dreamed in live</u> together.	[dreamed of living]
it should be <u>added in</u> its programming ...	[added to]

Notice that if we translate the errors regarding *in* as a dependent preposition into Portuguese, the preposition *em* is probably the one the subjects had in mind when they wrote these sentences. By referring to the contrast between the prepositional systems of both languages (Chapter 2), we find that *em* may have as English equivalents *on*, *at*, *about*, *to*, and *of*, which are the targets the subjects failed to use. Thus, on balance, this multiplicity of English prepositions for the Portuguese preposition *em* may be accounting for most of the errors shown above.

The other errors involving the preposition *in* were distributed throughout its use as dependent on nouns (XNPR) and adjectives (XADJPR), errors of complementation of nouns (XNCO), adjectives (XADJCO), and prepositions (XPRCO), overuse (WR), and underuse (WM). Below is a list of all of these errors (the error tag and the suggested correction are placed between brackets to the right of the examples):

they have to <u>pay attention in</u> the rhythm ...	[XNPR - pay attention to]
will always have an <u>influence in</u> their person ...	[XNPR - influence on]
a person has to have a great <u>knowledge in</u> his area ...	[XNPR - knowledge of]
the bad <u>influence</u> this standards might cause <u>in</u> their lives ...	[XNPR - influence on]
well-informed person, <u>updated in</u> what is happening ...	[XADJPR - updated on]
do not underestimate your <u>hability in learning</u> ...	[XNCO - ability to learn]
I really have a good <u>method in learning</u> ...	[XNCO - method to learn]
prove myself that I was <u>good enough in doing</u> it.	[XADJCO - good enough to do]
people who are interested <u>in get</u> more money ...	[XPRCO - in getting]
people are interested <u>in get</u> more and more money ...	[XPRCO - in getting]
government is not interested <u>in improve</u> educational ...	[XPRCO - in improving]
In spite of <u>to create</u> ...	[XPRCO - creating]
In spite of <u>to be</u> afraid ...	[XPRCO - being]

relationship at home, at work, and <u>in</u> everywhere ...	[WR - "0"]
<u>In</u> according to this researchers, ...	[WR - "0"]
<u>In</u> today, people are interested ...	[WR - "0"]
at least once <u>in</u> a week.	[WR - "0"]
understand what is going " <u>0</u> " the world ...	[WM - in]
we are living " <u>0</u> " a period of high unemployment ...	[WM - in]

Once again, if we seek the Portuguese equivalent preposition for the above sentences, the preposition *em* appears to be the one the subjects had in mind most of the time. Thus, even though the examples above extracted from the Br-ICLE bear witness to different grammatical errors, they do resemble each other in one aspect: the overgeneralization that the English preposition *in* may, most of the time, be substituted for the Portuguese preposition *em*, regardless of the semantic or grammatical functions the subjects want to convey.

All in all, in spite of its great frequency, the preposition *in* was wrongly employed only 8% of the time it was used. In many of the examples, the semantic contexts in which these errors occurred were related to place and time expressions. It has also been posited that in the great majority of cases, the target prepositions were either *at* or *on*, with fewer instances where *about*, *to*, and *of* seemed to be the correct alternatives. This fact is corroborated by the CA of both languages, which indicated that the Portuguese preposition *em* may have as English equivalents the prepositions mentioned above.

4.7 The preposition OF

Of was the most frequent preposition in the Br-ICLE accounting for 935 occurrences, 868 being regarded as correct (93%), and 67 as incorrect. Before describing the errors, how they were distributed within the error categories, and their likely sources, the following section will address the instances with which the subjects did not have problems.

4.1.7 Correct uses of OF

The syntactic uses and semantic relations conveyed by the preposition *of* were many and varied. This section is just an overview of the most common uses employed by the subjects of this study, not accounting for the entirety of correct uses of the preposition *of*.

The main uses of *of* regarded its use in nominal groups²⁴, when N₁ was a number or measurement of N₂, when N₁ offered support to N₂, in set phrases, when N₁ specified some part of N₂, and following some verb- and adjective-forms:

Semantic and syntactic uses	Number of correct uses	%
Nominal groups	220	25
N1 number or measurement of N2	183	21
Support	81	9
Set phrases	71	8
N1 as part of N2	51	6
Following verbs and adjectives	36	4
Others	226	27
TOTAL	868	100%

Table 13 – *Of* distribution of correct uses

Within nominal groups (NG), nominalizations, (i.e., propositional relationships between the two nouns), were one of the most common uses of the preposition *of*, as in the following examples:

	N ₁		N ₂
	this modification	of	thoughts
	the ruin	of	his life
	the values	of	television

In general we can say that the noun group allows for two nouns of equal status to be chosen and connected by *of*. In the examples above, neither noun seems to be dominant, and the structure simply requires both of them.

Of was also largely employed when N₁ was a number or a conventional measure of N₂. It can be noticed that N₂ is the most likely headword, for example:

N ₁		N ₂
hundreds	of	activities
a lot	of	money
a little bit	of	this

Here is a list of measure words which were used with *of* (frequencies in parenthesis):

all of (9) both of (5) a lot of (31) lots of (10) number of (19)
 some of (10) amount of (8) one of (24) majority of (6) most of (13)

Of was also commonly used when N₁ was seen as offering support to N₂, rather than only indicating a number or measures related to it, for example:

N ₁		N ₂
The concept	of	love
a good example	of	this
All kinds	of	cruelty

The expressions which were most used to convey this kind of support to N₂ were *concept of* (8), *example of* (10), *kind of* (48), *lack of* (11), and *sort of* (9).

Of was also used in structures where N₁ specified some part of N₂, for example:

N ₁		N ₂
the beginning	of	this century
The level	of	violence
the list	of	best-sellers

Of also appeared in a number of set phrases. The most common were *because of* (27), *first of all* (6), *instead of* (13), *of course* (16), and *in terms of* (6).

Finally, the preposition *of* appeared following certain verb- and adjective-forms (VA). Here is a list of some verbs and adjectives with which *of* was used as a dependent preposition:

Adjectives: *afraid, ashamed, aware, capable, conscious, free, full, proud, unaware, tired.*

Verbs: *die, dream, exclude, make, get rid, suppress, think, worship.*

According to Sinclair (1991, p.82-3) "Prepositions are mainly involved in combining with following nouns to produce prepositional phrases which function as adjuncts in clauses. This is not anything like the main role of *of*, which combines with preceding nouns to produce

²⁴ According to Sinclair (1991, p. 90), in nominal groups "neither noun seems to be pivotal or dominant".

elaborations of the nominal group”. Indeed, it has been shown in this section that most of the correct uses of *of* do involve elaborations of nominal groups, a feature which differentiates *of* from the other prepositions.

4.7.1 Incorrect use of OF

The errors regarding the preposition *of* were distributed within ten error categories.

Table 14²⁵ displays these error categories along with their frequency of errors:

Error categories	Number of incorrect uses	%
LP	19	2
DEP	14	2
CO	13	1
LS	5	8
WR	10	1
WM	6	9
TOTAL	67	100%

Table 14 – *Of*: distribution of errors

The errors involving *of* in fixed expressions or lexical phrases (LP) accounted for 28% of the total number of errors. Many of these instances cannot be regarded as *errors*, for the fact that they seem to provide examples of *non-native like* uses, rather than ungrammaticality. Below are all the instances regarding these misuses of *of* (the suggested correction is placed between brackets to the right of the example):

<u>the disappointment of Rita</u>	[Rita’s disappointment]
<u>the project of FHC’s reelection</u>	[the project to reelect FHC]
<u>the relationship of a couple</u>	[a couple’s relationship]
<u>the rhythm of each other</u>	[each other’s rhythm]
roots of increasing <u>rate of crimes</u>	[crime rate]
relationship with the increasing <u>rate of crimes</u>	[crime rate]
led them to increase the <u>rate of crimes</u>	[crime rate]
it would stablish the <u>rate of birth</u>	[birth rate]
in most societies, the <u>rate of crime</u>	[crime rate]
<u>the rate of born</u>	[birth rate]
we are the <u>opposite of</u> xenophobic	[opposite from]
when the <u>industry of media</u>	[media industry]
existing leaders are then, put <u>apart of</u> any leadership	[apart from]

²⁵ For the sake of simplicity I have lumped together the error categories (XVPR) and (XNPR) under the tag DEP, and the error categories (XADJCO), (XNCO), (XPRCO) and (XVCO) under the tag CO.

visit our houses through a <u>set of TV</u>	[TV set]
if you have an <u>opportunity of job</u>	[job opportunity]
have a <u>lace of pearls</u>	[pearl necklace]
the <u>race of the Jewish</u>	[Jewish race]
the brazilians, <u>carecent of</u> structure	[lacking in]
which should be <u>at the same level of</u> the other person	[at the same level as]

All but four of the examples above can be categorized into two types of infelicity. The first type is caused by a confusion as to when to use genitives to convey possession. Because of equivalences like *the car of John*, and *John's car* it is said that the N_1 of N_2 is an alternative way of stating that N_2 possesses N_1 . In fact, the *of* structure may have little to do with ownership or possession, as can be seen when a personal pronoun in N_2 position has to be expressed in the possessive form, for example, *a friend of mine*, not *a friend of me*.

The second type of infelicity refers to *noun + noun* structures which are usually seen as fixed expressions. Notice that the noun *rate* was misused with the nouns *birth* and *crime* in fixed expressions such as *birth rate* and *crime rate*.

The problem with these two types of infelicities seems to be related to a lack of knowledge of *noun + noun* structures as well as a lack of knowledge of pseudo-possessive structures.

Errors involving the independent use of *of* (LS) accounted for 8% of the total number of errors. Below are all the instances regarding these misuses of *of* (the suggested correction is placed between brackets to the right of the example):

they learn behaviors from the programs <u>of</u> television	[on]
For instance, the students <u>of</u> the Course in Mathematics	[from, in]
but your diploma is <u>of</u> engineering	[in]
I am a student <u>of</u> the Course in Letras	[from, in]
Suharto finally stepped down after a pression <u>of</u> students	[from]

The explanation for the errors above may be sought through the translation of these sentences into Portuguese. In most of the examples, the Portuguese preposition *de* appears to be what the subjects had in mind when they wrote these sentences. It has been suggested in

Chapter 2 that *de* might have as English counterparts the prepositions *of* and *from*. As can be noticed in three of the five examples the target preposition would most likely be *from*, the two exceptions being *on television* and *in engineering*. Thus, it may be argued that once again the multiplicity of forms of English prepositions in relation to Portuguese prepositions may be contributing to the production of such infelicities. Furthermore, the subjects seemed to have overgeneralized the preposition *of* as the only equivalent of *de*.

Overuse errors (WR) accounted for 15% of the total number of errors. These errors indicate that the subjects overused *of* with the adverb *most*, with the connectors *besides* and *despite*, and with four other sentences (the tag [0] indicates that *of* should have been omitted):

most <u>of</u> people get very depressed ...	[0]
most <u>of</u> people doesn't share the same...	[0]
most <u>of</u> people spend hours of their lives ...	[0]
Besides <u>of</u> stimulating violence ...	[0]
besides <u>of</u> giving you the knowledge ...	[0]
Despite <u>of</u> receiving bad news ...	[0]
to sleep inside <u>of</u> your companion ...	[0]
people have the right to give up <u>of</u> the life ...	[0]
people that do not mind <u>of</u> being young forever ...	[0]
these laws lack <u>of</u> basic human rights ...	[0]

In Portuguese we say *a maioria da(s)/do(s)*, *além de*, and *apesar de*, which are structured with the preposition *de*, whereas in English we use *most people = a maioria das pessoas*, *besides = além de*, and *despite = apesar de*. For this reason, these errors may not come as totally unexpected, and seem to stem from L1 transfer. The sentence *sleep inside of your companion* may be viewed as an odd way of saying 'to have sex', and it is very unlikely that a native speaker would produce such a sentence. As to *give up of the life* meaning *desistir da vida*, it appears as though the subject who produced this sentence was searching for a literal translation where *desistir = give up*, *da = of the*, and *vida = life*. Translating the last two errors, *do not mind of being* and *laws lack of basic human rights*, into Portuguese, *não se*

importam de ser and *leis tem falta de direitos humanos*, reveals that the subjects who produced these sentences were also searching for an equivalent of *de* in English, and for this reason, they might have used the English preposition *of*. Thus, it may be argued that these last errors may, as well, stem from L1 transfer.

Of was also misused as dependent on nouns (XNPR) and verbs (XVPR) in 21% of the errors (DEP). Here is a list of the errors involving these two error categories (the suggested correction is placed between brackets to the right of the example):

The <u>synonin of</u> having a privileged life ...	[XNPR – synonym for]
the <u>place of</u> evaluation is inside your mind ...	[XNPR – place for]
love is not the <u>desire of</u> creating a masterpiece ...	[XNPR – desire for]
it just means a plundering <u>desire of</u> life ...	[XNPR – desire for]
having a new <u>perspective of</u> life inside society ...	[XNPR – perspective on]
provoke a <u>decrease of</u> the world natural sources ...	[XNPR – decrease in]
another example is the <u>course of</u> digital design ...	[XNPR – course in]
humanity do not pay <u>heed of</u> it ...	[XNPR – heed to]
the <u>incentive of</u> her parents and <u>of</u> her husband ...	[XNPR – incentive from]
He was satisfied and <u>filled of</u> victory ...	[XVPR – filled with]
life inside society and not <u>excluded of</u> it ...	[XVPR – excluded from]
it <u>depends of</u> the abilities of each one ...	[XVPR – depends on]
it <u>depends of</u> human spiritual evolution ...	[XVPR – depends on]
<u>separate</u> political unsolved problem <u>of</u> real unsolved problems ...	[XVPR – separate from]

Noun and verb dependent prepositions can be said to have a very idiosyncratic patterning, which in turn results in great difficulties for EFL learners. It seems that in the errors listed above, the preposition appears as a delexicalized item, most of the meaning being conveyed by the nouns and verbs which precede it. Translating the above nouns and verbs into Portuguese reveals that all of them may combine with the Portuguese preposition *de*; as in *sinônimo de*, *lugar de*, *perspectiva de*, *desejo de*, *depende de*, etc.; which may be one of the reasons, besides the irregular patterning of dependent prepositions, for the errors displayed above.

Another type of error involved the wrong complementation of the preposition *of* (XPRCO), and its wrong complementation with nouns (XNCO) verbs (XVCO), and

adjectives (XADJCO), all of those collapsed into the category CO. This category accounted for 19% of the total number of errors (the tags and suggested corrections are placed between brackets to the right of the example):

he was afraid <u>of fail</u> ...	[XADJCO - afraid of failing]
regret the fact <u>of not have</u> known	[XPRCO - of not having known]
<u>Instead of use</u> glasses ...	[XPRCO - instead of using]
the <u>ability of painting</u> ...	[XNCO – ability to paint]
the <u>capacity of making</u> ...	[XNCO – capacity to make]
respect and <u>desire of learning</u> ...	[XNCO – desire to learn]
was the <u>desire of learning</u> ...	[XNCO – desire to learn]
the <u>desire of being</u> together ...	[XNCO – desire to be]
the <u>need of coming</u> back ...	[XNCO – need to come]
not having the <u>rights of claiming</u> ...	[XNCO – right to claim]
it's a strange <u>sensation of thinking</u> that two people ...	[XNCO – sensation to think]
wait for the <u>right time of doing</u> it ...	[XNCO – right time to do]
there is <u>urge of keeping</u> some values ...	[XNCO – urge to]

In the errors above, the subjects either failed to use a gerund verb after *of*, or used the gerund form when they were supposed to use infinitive. One possible explanation for not using the gerund form may be sought in their Portuguese grammar: *medo de falhar*; *arrepender-se do fato de não ter sabido*; *ao invés de usar*; where it seems that a *de+infinitive* structure is being generalized for an *of+infinitive* structure. However, the opposite also happened, the subjects made use of gerund verbs when they should have used infinitive ones. Thus, it is difficult to ascertain the source(s) of these errors since an appeal to the subject's natural grammar does not seem to suffice us with convincing evidence for what is taking place.

Two other category, accounting for 9% of the errors, regarded the omission (WM) of the preposition *of* (I have inserted the tag 0 in the examples to indicate where the preposition *of* has been omitted):

Being very young is something very hard to think <u>0</u> .	[of]
take care <u>0</u> them ...	[of]
we see a lot of fishes dying because <u>0</u> contamination ...	[of]
Due to all <u>0</u> this, in my opinion, ...	[of]

Thousands <u>0</u> years ago, ...	[of]
she has changed all her life because <u>0</u> love ...	[of]

The preposition *of* was correctly employed 93% of the time it was used. Most of its correct use regarded nominalizations, and instances where N_1 was a measure of N_2 . This fact indicates that *of* was used in a very different way compared to the other prepositions, which usually combined with following nouns to produce prepositional phrases functioning as adjuncts in clauses.

Most of the errors were related to problems with genitives and compound nouns, misuse of *of* as dependent on nouns and verbs, and the complementation of nouns, verbs, and adjectives. By comparing both languages, it has been suggested that the subjects' L1 grammar, especially the use of the preposition *de* to construct possessives, as in *a casa de meu pai*, compound nouns, as in *taxa de crime*, as a preposition dependent on nouns and verbs, for example, *perspectiva de reabilitação*; *depende da habilidade de cada um*; and as the complement of adjectives, nouns, prepositions, and verbs, as in *medo de falhar*; *capacidade de pintar*; *ao invés de*; *pessoas não se importam de ser*; might be one of the factors which has contributed to the misuses of the preposition *of*.

4.8 The preposition ON

On accounted for 164 occurrences in the Br-ICLE, 138 being regarded as correct (84%), and 26 as incorrect. Before describing the errors, how they were distributed within the error categories, and their likely sources, the following section will address the instances with which the subjects did not have problems.

4.8.1 Correct uses of ON

The correct uses of the preposition *on* were distributed among place expressions, verb dependent uses, connectors, phrasal verbs, time expressions, and other uses. Table 15 displays the distribution of correct uses regarding the preposition *on*:

Semantic and syntactic uses	Number of correct uses	%
Place	37	27
Verbs	36	26
Part of connectors	29	21
Phrasal verbs	12	9
Time	7	5
Others	17	12
TOTAL	138	100%

Table 15 – On: distribution of correct uses

Place expressions accounted for 27% of the correct uses of *on*. The most common places indicated by *on* were *on TV* (13), *on the street(s)* (5), *on earth* (5), and others like *on stage*, *on the sofa*, *on Wall Street*, etc.

On was employed as dependent on verbs in 26% of its correct uses. Some of the verbs occurring with *on* were: *base* (11); *concentrate* (2); *cheat* (1); *depend* (8); *focus* (2); *impose* (3); and *take part* (1).

Regarding its use as part of connectors, which accounted for 21% of the correct uses, *on* was employed to indicate contrast in connectors like *on the other hand* (12) and *on the contrary* (4), to explicate something as in *on the grounds that* (3), and to exemplify something as in *and so on* (10).

On was also used as a particle of the phrasal verbs *go on* (6), *look down on* (1), *cut down on* (1), *keep on* (1), and *turn on* (3). This use of *on* accounted for 9% of its correct use.

On was employed 5% of the time to convey time relations in expressions like *on the eve*, *later on*, *on weekends*, *on vacation*, etc.

The other uses of *on* accounted for 12% of its correct use, and were distributed among adjective and noun dependent uses, as in *outlook on life*; and fixed expressions such as *on the verge* and *on the pill*.

4.8.2 Incorrect uses of ON

The errors involving the preposition *on* were distributed within its independent use (LS), its dependent use on verbs (XVPR), nouns (XNPR), and adjectives (XADJPR), and its overuse (WR), as it can be seen in Table 16:

Error categories	Number of incorrect uses	%
LS	10	3
XVPR	6	2
XNPR	5	1
XADJPR	2	8
WR	3	1
TOTAL	26	100%

Table 16 – *On*: distribution of correct uses

Independent uses of the preposition *on* (LS) accounted for 38% of the total number of errors. The list of errors below indicates that *on* was misused as an independent preposition in two semantic contexts: place, sometimes metaphorically or figuratively, and once to indicate time. In most of the errors the target preposition was *in*, followed by *at*, *with*, and *from*, which were also the target in fewer cases (the suggested correction is placed between brackets to the right of the example):

plastic modelling increased from 8% to 90% <u>on</u> the last 5 years ...	[in]
walking <u>on</u> the park or zoo ...	[in, at]
they are finally <u>on</u> the right direction ...	[in]
<u>On</u> countries that apply limitation laws ...	[in]
the solution is <u>on</u> the government's hands ...	[in]
a large number of people competing <u>on</u> the job market ...	[in]
only by working <u>on</u> a certain field ...	[in]
you picked it up <u>on</u> the nearest grocery ...	[at]
possible to see our prejudice <u>on</u> the outside ...	[from]
became an incentive to her to continue <u>on</u> her project ...	[with]

The errors above bear witness to the confusion the subjects make when using *in*, *on*, and *at* to convey both spatial and temporal relations. In most of the above errors, *em* was the equivalent preposition in Portuguese. As with the errors involving independent uses of *at* and

in, once again the errors involving independent uses of *on* appear to stem from the multiplicity of forms the preposition *em* can have in English.

On was wrongly employed as dependent on verbs (XVPR) in 23% of the errors. Except for the first two, *culpado por/de* and *dedicar-se à* as the corresponding Portuguese expressions, all errors contain verbs which would probably take the Portuguese preposition *em*: *interferir em*; *viver em*; *levar em consideração*; *prestar atenção em*. Thus, at least for these latter errors, the subjects' L1 might be influencing their prepositional choice. In addition, if most of these verbs take *em* as a dependent preposition in Portuguese, it may be expected that errors like the ones above occur.

they have no fear of being <u>blamed on</u> anything ...	[blamed for]
they have to <u>devote</u> themselves <u>on</u> this plot ...	[devote to]
has no right to <u>interfere on</u> people's personal affairs ...	[interfere in]
we <u>live on</u> a constant fight ...	[live in]
if we <u>take on</u> account that ...	[take into account]
you should <u>pay attention on</u> the words you use ...	[pay attention to]

On was erroneously used as dependent on nouns (XNPR) in 19% of the errors. Regarding the first error, the structure *bringing confusion on dealing* is totally ungrammatical and unintelligible. As for the second sentence, the subject who made this error was probably not aware of the fact that in English the noun *purpose* requires the preposition *in* instead of *on*, confusing *propósito em* with *purpose on*. In the last three errors *on* was used as an equivalent of the Portuguese prepositions *para*, *à*, and *sobre*.

bringing <u>confusion on</u> dealing with several steps ...	[confusion to]
Is there any <u>purpose on</u> it?	[purpose in]
There is no <u>recipy on</u> how to love ...	[recipy for]
having no <u>rights on</u> his money ...	[rights to]
they have <u>knowledge on</u> many areas ...	[knowledge of]

Misuses of *on* as dependent on adjectives accounted for 8% of its errors. Once again *on* was generalized as the equivalent to the Portuguese preposition *em* regardless of the

adjectives that preceded it. Notice that in Portuguese we say *interessado em*, and *bem sucedido em*, and for this fact, it is not surprising that the subjects made the errors below.

the teacher was not so <u>interested on</u> teaching ...	[interested in]
to be <u>successful on</u> their searches ...	[successful in]

The last errors involving the preposition *on* were related to its overuse (WR), and accounted for 12% of its errors (the tag “0” is placed to the right of the example to indicate that *on* should not have been used). It can be seen that, in Portuguese, the verb *believe* requires the preposition *em*, as in *Eu acreditei no depoimento dele*, whereas in English, we may say *I believed his report*, which, in turn, may be a possible reason for these two errors. The other error, *said on to the other*, might reflect just a slip of the pen, being difficult to assert other possible reasons.

they believe <u>on</u> what they see ...	["0"]
the person said <u>on</u> to the other who he had fooled ...	["0"]
most people must believe <u>on</u> it ...	["0"]

Overall, the preposition *on* was correctly used in three major contexts: to indicate place, as dependent on verbs, and as part of different connectors. The errors were concentrated in three categories: (LS), (XVPR), and (XNPR). It has been argued that most of the errors probably derive from the fact that the Portuguese preposition *em*, which was in most cases the preposition the subjects would use in their mother tongue, has many equivalents in English (e.g., *at, in, on, about*) and that this multiplicity of forms may be a factor of confusion for the subjects.

4.9 The preposition TO

The preposition *to* accounted for 280 occurrences in the Br-ICLE, 229 occurrences (82%) being regarded as correct and 51 occurrences as incorrect. Before describing the most frequent errors, their categories, and their likely sources, I will point out some of the uses with which the subjects did not have problems.

4.9.1 Correct uses of TO

In most of its correct uses, the preposition *to* was syntactically used as dependent on verbs, nouns, and adjectives, as well as to form complex prepositions. Table 17 displays the frequency of these syntactic uses:

Semantic and syntactic uses	Number of correct uses	%
Verbs	97	4
Nouns	34	1
Complex prepositions	31	1
Adjectives	17	6
Others	50	2
TOTAL	229	100%

Table 17 – To: distribution of correct uses

Table 17 indicates that *to* was used as dependent on verbs, nouns, and adjectives in 63% of its correct use. Below I show some of the verbs, nouns, and adjectives which preceded *to* (all the verbs are in the stem form):

Verbs

apply (2) belong (3) come (3) connect (2) contribute (4) do (5) give (4)
happen (3) listen (3) need (2) relate (9) talk (6) send (4) travel (3)

Nouns

access (2) incentive (2) solution (4) burden (2) apology (1) warning (1)

Adjectives

essential (3) important (5) equal (3) vulnerable (1) unfair (1) passive (1)
relevant (1) integrated (1)

The complex prepositions *according to* and *due to* accounted for 14% of the correct uses of the preposition *to*. The other occurrences of *to* were related to its use as part of connectors (e.g., *in addition to*; *when it comes to*; *in comparison to*; *in relation to*), to occurrences where *to* was used to start a sentence, as in *To scientists, these numbers are a consequence of ...*; and to express personal opinion: *It seems to me that these almost 20 years of democracy* Altogether, these other uses of *to* accounted for 20% of its correct use.

4.9.2 Incorrect uses of TO

The errors involving the preposition *to* were distributed among eight error categories. Three error categories, namely, (XVPR), (XNPR), and (XADJPR) regarded its misuse as dependent on verbs, nouns, and adjectives. Three other error categories, namely, (XVCO), (XNCO), and (XPRCO), regarded its misuse as the complement of verbs, nouns and prepositions. The other two error categories, regarded its misuse as an independent preposition (LS), and its avoidance (WM). Table 18 displays the frequency of errors among these eight of error categories:

Error categories	Number of incorrect uses	%
LS	12	2
XVPR	11	2
XNPR	11	2
XADJCO	1	2
XVCO	7	1
XNCO	2	4
XPRCO	3	6
WM	4	8
TOTAL	51	100%

Table 18 – *To*: distribution of errors

As can be seen in Table 18, the error categories (LS), (XVPR), and (XNPR), accounted for most of the errors involving the preposition *to*. The errors involving independent uses of *to* (LS) are listed below (the suggested correction is placed between brackets to the right of the example):

observe that new devices <u>to</u> home - vacuum cleaner, micro waves ...	[for]
Hard times <u>to</u> them.	[for]
So, some wanted to succeeded on their own, <u>to</u> their own satisfaction ...	[for]
marriage was for interest or just <u>to</u> the women become a housewife ...	[for]
But the mainly importance <u>to</u> everything happens is the love ...	[?for]
women were just tools <u>to</u> nobles become rich ...	[for]
the expectancy <u>to</u> 2005 was about 7 billions.	[for]
"the doors" are always open <u>to</u> him ...	[for]
and eventually organs lost <u>to</u> disease.	[because of]
Exercises should be repeated and checked <u>to</u> the original pattern ...	[with]
could bring some pleasure <u>to</u> your work.	[into]
they clearly manipulate the reader's ideas <u>to</u> their point of view ...	[according to]

In most of the semantic contexts where the errors above appeared, the preposition *to* was used in dative expressions (e.g., *the doors are always opened to him; women were just tools to nobles*). With the exception of *manipulate reader's ideas to their point of view...*, and *Exercises should be repeated and checked to the original pattern ...*, and *... the mainly importance to everything happens* (which is ungrammatical and unintelligible), in all of the errors the Portuguese preposition the subjects probably had in mind when they wrote the sentences was *para*. It has been shown in Chapter 2 that the Portuguese preposition *para* may have as English equivalents both, *to* and *for*. In fact, as most of the errors attest, *for* was the target preposition the subjects failed to employ. Thus, for the errors above, it might be plausible to argue that L1 interference may be accounting for the misuses of the preposition *to*.

The errors regarding *to* as dependent on verbs (XVPR) are displayed below (the suggested correction is placed between brackets to the right of the example):

<u>advertising</u> us <u>to</u> the danger of few in command ...	[warning us about]
it was also <u>associated to</u> it ...	[associated with]
money which is <u>destinated to</u> something ...	[destined for]
Clothes stores <u>direction</u> their products <u>to</u> the public ...	[direct towards]
As a result, they do not <u>interact to</u> each other ...	[interact with]
If we <u>look to</u> the art or <u>to</u> the love as something ...	[look at]
please turn off the television set and <u>look to</u> your family ...	[look at]
theory is not enough to <u>prepare</u> professionals <u>to</u> the job market ...	[prepare for]
have more capacities and to <u>prepare to</u> important business ...	[prepare for]
she wants to <u>prepare to</u> her professional life.	[prepare for]
to <u>provide</u> good education, health care and other services <u>to</u> ...	[provide for]

In relation to the errors which had as target prepositions *with* and *about*, the explanation may be related to the fact that verb-dependent prepositions behave in a very idiosyncratic way. Nonetheless, if we translate into Portuguese these errors: *advertir sobre*, *concordar com*, *associado com*, and *interagir com*, it is not clear why the subjects failed to employ the prepositions *with* and *about*. The same may be suggested about the errors that required *for*, *at*, and *towards* as target prepositions: *destinar para*, *preparar para*, *prover para*, *olhar para*,

and *direcionar para* (especially because the subjects did not use *for*, which is the most likely translation for *para*). These examples seem to demonstrate that learners do not always fall back on their L1 when they lack knowledge about the L2.

The errors below regard misuses of *to* as dependent on nouns (XNPR). As can be noticed, with the exception of *dangerous issues to brinkmanship*, which required *on* as the target preposition, all of the other errors required the preposition *for*. This fact seems to reinforce what has been suggested in the first error category: L1 interference seems to be the major cause for these errors. Translated into Portuguese, all of the nouns below require the preposition *para*, which has as possible English counterparts *to* and *for*.

holds up standards for personal <u>accomplishment to</u> children ...	[accomplishment for]
the <u>consequences of</u> this increase <u>to</u> the world ...	[consequences for]
procrastinate dangerous <u>issues to</u> brinkmanship ...	[issues on]
university opens the <u>opportunities to</u> many people ...	[opportunities for]
Does university open <u>opportunities to</u> the market?	[opportunities for]
Great <u>opportunities to</u> everyone ...	[opportunities for]
there is no <u>place to</u> man's feelings ...	[place for]
there are severe <u>punishments to</u> the practitioners of a crime ...	[punishment for]
The government should take the <u>responsibility to</u> this ...	[responsibility for]
Teenagers are the main <u>target to</u> companies advertisements.	[target for]
If television is a source of <u>values to</u> young people ...	[values for]

Errors of complementation encompassed misuses with adjectives (XADJCO), verbs (XVCO), nouns (XNCO), and infinitive uses after prepositions (XPRCO), and accounted for 26 per cent of the errors involving the preposition *to*. Below is the complete list of complementation errors (the error category tag and the suggested correction are placed between brackets to the right of the example):

things started changing and becoming <u>easier to</u> the pigs.	[XADJCO - easier for]
we could <u>ask to</u> ourselves - How do they do that?	[XVCO - ask ourselves]
philosophy would <u>bring to</u> the people the opportunity to live	[XVCO - give people]
the emotional fact, which <u>concerns to</u> the family ...	[XVCO - concerns the family]
each prisoner <u>costs to</u> our government ...	[XVCO - costs our government]
<u>offering to</u> the student richer courses ...	[XVCO - offering the student]
they <u>showed to</u> everyone that they can do ...	[XVCO - showed everyone]
<u>showing to</u> the people only what they want ...	[XVCO - showed people]
the author's creative <u>way to</u> tell the Russian revolution ...	[XNCO - way of telling]
different <u>ways to</u> do something ...	[XNCO - ways of doing]

<u>After to feel</u> accepted a feeling of panic ...	[XPRCO - after feeling]
think a lot <u>before to take</u> a decision.	[XPRCO - before taking]
think <u>before to say</u> what they really want ...	[XPRCO - before saying]

In each of the four error categories above, the subjects appear to have employed a different strategy. In relation to the wrong complementation of the comparative adjective *easier*, the dative relation was expressed with *to* instead of *for*, and might stem from the fact that we say *mais fácil para você* in Portuguese (notice again the confusion between *to* and *for* in dative relations).

As for the wrong complementation of verbs, the subjects seem to have disregarded the fact that in English these verbs are usually constructed as *verb + object (indirect) + object (direct)*, and that they do not require the preposition *to* in the indirect object. Notice as well, that in all of these errors the subjects were trying to convey dative relations (e.g., **ask to ourselves*; **showed to everyone*), and that in Brazilian Portuguese, contrary to English, the verbs would be constructed using the preposition *para* (e.g., *perguntar para nós mesmos*; *mostrar para todo mundo*). Thus, it seems that L1 interference may be a likely cause for the errors involving the complementation of verbs.

The noun *way*, meaning a particular manner or style of behavior, was wrongly complemented two times. It is possible to say in English *I'm on my way to school*, but it is very unlikely that the sentence *?A creative way to tell the story* would be uttered by a native speaker. In addition, in Portuguese we may say *Uma maneira criativa de contar a história*, where the verb appears in the infinitive form, whereas in English we would say *A creative way of telling the story*, the verb being in the gerund form. Therefore, there appears to exist a relationship between using the infinitive form after the noun *maneira*, in Portuguese, and attempting to do the same with the noun *way*, in English.

The last errors of complementation involved the prepositions *before* and *after*. As argued in the above paragraph, the errors involving these two prepositions seem to reflect

nearly the same case: the fact that in Portuguese we say *antes de dizer*, where we use the complex preposition *antes de* and the verb *dizer* in the infinitive form, whereas in English we say *before saying*, using a simple preposition and the verb in the gerund form.

Avoidance errors (WM) constituted the last error category, accounting for eight per cent of the errors regarding *to*. Since it is not clear (at least to me) whether these errors may be explained as slips of the pen or not, I will not put forward a theory to try to explain them (the tag “0” indicates the place in the sentence where *to* has been omitted):

according “0” magazine "Veja"(n. 15, 14th April 99, pag.81)	[to]
they are not paying attention “0” what that person on TV is saying.	[to]
According “0” the academic standards ...	[to]
The cells which give rise “0” the more different tissues in our body.	[to]

In sum, the error categories (LS), (XVPR), (XNPR), and (XVCO) accounted for 80 per cent of the errors involving the preposition *to*. Out of the fifty-one errors analyzed, twenty-four required as the target preposition the preposition *for*. It has been suggested that most of the errors, even those which did not require *for* as a target preposition, seem to stem from L1 interference or from the fact that the Portuguese preposition *para* may have as English equivalents *for*, *to*, *at*, and *towards*.

4.10 The preposition WITH

The preposition *with* accounted for 167 occurrences in the Br-ICLE, 155 occurrences (93%) being regarded as correct and 12 occurrences as incorrect. Before describing the most frequent errors, their categories, and their likely sources, I will point out some of the uses with which the subjects did not have problems.

4.10.2 Correct uses of WITH

The preposition *with* was employed to convey five semantic relations: accompaniment, the notion of ‘having’, instrument, manner, and support. Table 19 displays the distribution of frequencies for each of the above semantic relations:

Semantic and syntactic uses	Number of correct uses	%
Accompaniment	35	2
Having'	28	1
Instrument	6	4
Manner	33	2
Support	36	2
Others	17	1
TOTAL	155	100%

Table 19 – *With*: distribution of correct uses

The notion of accompaniment, that is, the meaning ‘in company with’ or ‘together with’, was specially followed by animate complements, as in *It’s not so terrible to live with somebody* or *Women are sharing their bills with their husbands*.

The idea of possession, or ‘having’, appeared in sentences like *...while children with siblings are not supported by the government...* or *A person with blond hair, blue eyes and button nose...*

With was employed to express the meaning of instrument in sentences such as *There are plenty of books with the aim of teaching the reader...* or *...and only have children when they can support them with food, education, ...*

Manner was expressed through *with* in sentences like *Everybody has the right to live with dignity and quality* or *...TV rushes into this opening with a world packed in living colors*.

Finally, the notion of support, that is, expressions of solidarity or sympathy, was also conveyed by *with* in sentences such as *In a way I agree with Greenpeace and other organizations...* or *...the Nazi idealism is connected with some standards such as blond angels...*

4.10.3 Incorrect uses of WITH

The errors involving the preposition *with* were greatly concentrated in its use as dependent on verbs (XVPR). Two other error categories, namely, (LS) and (XADJPR), accounted for the rest of the errors regarding the preposition *with* as indicated in Table 20:

Error categories	Number of incorrect uses	%
XVPR	9	7
LS	2	1
XADJPR	1	8
TOTAL	12	100%

Table 20 – *With*: distribution of errors

As can be seen in Table 4.18, three fourths of the errors regarded the use of *with* as dependent on verbs. The examples below appear to indicate that L1 interference seems to be playing a role in these errors, since in most of the sentences the Portuguese preposition *com* appears to be what the subjects had in mind. The English preposition *with* semantically bears a one-to-one correspondence with the Portuguese preposition *com*, and by translating the verbs into Portuguese (e.g. *aprender com*, *casar com*, *sofrer com*), we notice that the subjects disregarded or were not aware of the fact that these verbs, at least in English, do not take *with* as their dependent preposition, employing the preposition which they would use in Portuguese: *com* (the suggested correction is placed between brackets to the right of the example).

to <u>contribute with</u> their future ...	[contribute to]
What will <u>happen with</u> them?	[happen to]
It's normal and can <u>happens with</u> everybody.	[happen to]
but still <u>happening with</u> the ideas of our society ...	[happening to]
what they had <u>learned with</u> it	[learned from]
to know the person you will get <u>married with</u> as well ...	[married to]
the one <u>related with</u> the fact that all politicians ...	[related to]
we will not <u>suffer with</u> a terrible illness ...	[suffer from]
This new code of laws <u>works with</u> the idea that ...	[works on]

Independent misuses (LS) of the preposition *with* accounted for 17 per cent of its errors. In at least one of the examples below, once again it appears as though L1 interference seems to be affecting the prepositional choice in the sentence: *casamento com igreja, vestido, arroz*. As for the other example, it is quite difficult to infer its source. It can be deduced that the meaning the subject wanted to convey was 'money under German control', the Germans being represented by the noun **suastic*.

Therefore the tradition of marriage with church, wedding dress, rice ... [in]
 money in gold deposited in swiss banks with the suastic is incredible ... [under]

The last error regarded the misuse of *with* as dependent on adjectives (XADJPR). This last error, for which it is difficult to assert the sources, does not appear to derive from L1 interference since in Portuguese we would say: *Estas mulheres fazem um trabalho artístico similar ao do poeta.*

These women do a similar artistic work with the poet ... [similar to]

On balance, then, most of the errors involving the preposition *with* were related to its use as dependent on verbs and seem to stem from L1 interference, since in most of the examples the verbs used with the preposition *with*, would require the preposition *com* in Portuguese, whereas in English the subjects should have used other prepositions (e.g., *to*, *from*).

The following section addresses the research questions presented in chapter 3. The answers to these questions are based on what has been described from section 4.1 to 4.10, where an account of all the errors of each preposition together with their most common patterns of correct use has been put forward.

4.11 Addressing the research questions

In this section, each research question will be addressed and answered separately. Most of the evidence provided to support the answers derives from sections 4.1 to 4.10.

Research question 1: Is error concentrated on a particular preposition?

In order to answer this question, it is necessary to point out that the criterion for selecting the prepositions which were most misused was a ratio between the total number of tokens and the number of errors. Thus, the prepositions which yielded the highest number of errors were not necessarily the ones which were most frequently misused. For example, the preposition *of* occurred 935 times in the Br-ICLE accounting for 67 errors, more than any

type of preposition. However, the ratio between the total number of tokens and the number of errors was only 7 per cent, which means that *of* was correctly used 93 per cent of the time it was employed.

Table 21 shows the distribution of correct and incorrect uses of each preposition, as well as their error rates. It can be seen that the prepositions *to*, *at*, *by*, and *on* were the ones with which the subjects had the most problems. The preposition *about* was misused twelve per cent of the time it was employed, and the other prepositions, *for*, *from*, *in*, *of*, and *with*, were all misused less than ten per cent of the time they were used.

Thus, the answer to the question presented above is *no*. Error was not concentrated on *one* particular preposition but on *four* prepositions, which together accounted for 39 per cent of the total number of errors found in the present study (111 errors).

Preposition	Tokens	Correct use	Incorrect use	Error rate
About	112	99	13	12
At	84	70	14	17
By	119	99	20	17
For	268	251	17	6
From	84	80	4	5
In	717	658	59	8
Of	935	868	67	7
On	164	138	26	16
To	280	229	51	18
With	167	155	12	7
Total	2930	2647	283	100%

Table 21 – Error rate distribution

Research question 2: Is error concentrated in a particular error category?

Four error categories, namely, (LS), (LP), (XVPR), and (XNPR), accounted for 69 per cent of the total number of errors. As can be seen in Table 4, independent uses (LS), and verb-dependent uses (XVPR) of these prepositions accounted for nearly half of the total number of errors (44 per cent). The prepositions which accounted for most of the (LS) errors were: *in* (18), *on* (10), *to* (12), and *at* (8), and the prepositions which accounted for most (XVPR) errors were *to* (12), *with* (9), *in* (11), and *by* (10).

Error Category	No. of errors	%
LS	67	24
LP	34	12
XVPR	58	20
XNPR	36	13
XADJPR	10	4
XVCO	9	3
XNCO	18	6
XADJCO	3	1
XPRCO	11	4
WM	14	5
WR	23	8
Total	283	100%

Table 22 – Distribution of errors among various error categories

Research question 3: Which semantic/syntactic use is the most problematic for each one of the ten prepositions under investigation?

- a) **ABOUT:** In most of the errors involving this preposition (71 per cent), the semantic meaning ‘on the subject of’, ‘concerning’, was wrongly employed with nouns, adjectives, and verbs. The subjects appear to have generalized *about* as the most likely equivalent English preposition for the Portuguese preposition *sobre*, disregarding or not knowing the fact that some nouns, adjectives, and verbs do not collocate with *about* in English (see section 4.2).
- b) **AT:** Place and time expressions accounted for 79 per cent of the total number of errors involving this preposition. Both time and place misuses of *at* seem to indicate that the subjects have not made a distinction between *at* and *in* regarding a place as a point, without any real size, and to indicate the very specific time something is happening. Instead, the subjects appear to have generalized *in* as the preposition to convey these meanings.
- c) **BY:** In most of the errors regarding this preposition (74 per cent), the semantic contexts were related to either ‘agent’ or ‘cause’. Since the preposition *by* is chiefly used with

past participle verbs and some past participle adjectives, it appears as though, in many errors, the subjects erroneously employed it when making use of this verb form.

- d) **FOR:** The preposition *for* was erroneously used as dependent on adjectives, nouns and verbs in 41 per cent of the errors, and it was misused as the complement of adjectives, nouns, and verbs in 29 per cent of the errors. The semantic contexts in which these syntactic errors appeared were varied, not revealing a particular context with which the subjects seemed to have more difficulties.
- e) **FROM:** Only four errors were found regarding this particular preposition, two of them being related to 'source', and the other two being related to 'origin'.
- f) **IN:** Out of the fifty-nine errors involving *in*, twenty (34 per cent) were related to place expressions, and eight (14 per cent) were related to time expressions. In most of these contexts, the target prepositions were either *at*, or *on*.
- g) **OF:** Errors involving *of* in fixed expressions or lexical phrases accounted for 31% of the total number of errors. Many of these instances cannot be regarded as *errors*, for the fact that they seem to provide examples of *non-native-like* uses, rather than ungrammaticality. Most of these infelicities were related to a confusion with *noun + noun* structures which are usually seen as fixed expressions, and between using or not pseudo-genitives to convey possession.
- h) **ON:** Place expressions accounted for thirty-eight per cent of the errors involving *on* (10 errors). In most of these contexts the target prepositions were either *at*, or *in*.
- i) **TO:** The semantic contexts in which most of the errors (65 per cent) occurred were either when *to* was used to indicate who or what received something or had an action or feeling directed toward (e.g., **They don't give credit to the ones of real importance...*), or when *to* was used with some words to indicate what something is connected to or compared with (e.g., *... *money which is destinated to something...*). In 44 per cent of

the errors (24 occurrences), the English preposition *for* was the target preposition the subjects failed to supply.

- j) **WITH:** Errors involving the preposition *with* were chiefly concentrated in its dependent use on verbs (75 per cent). In 68 per cent of the errors (8 occurrences) the target preposition the subjects failed to employ was *to*. The contexts in which the errors occurred were varied, making it difficult to reveal semantic patterns of misuse.

Research question 4: Are the spatial and temporal uses of *in*, *on*, and *at* still problematic for the subjects even though they are assumed to have acquired these structures earlier on in their learning process?

Altogether, *in*, *on*, and *at* accounted for 99 errors (35 per cent of the total number of errors). Out of these 99 errors, 49 per cent of them were related to spatial and temporal expressions as indicated in Table 23:

Preposition	Place	Time	Others	Total
In	20	8	31	59
On	9	1	16	26
At	7	4	3	14
Total	36	13	50	99

Table 23 – *In*, *on*, and *at*: distribution of spatial and temporal errors

Spatial and temporal expressions involving these three prepositions accounted for 17 per cent of the total number of errors found in the present study. This very fact indicates that even though spatial and temporal expressions involving these three prepositions are among the first ones to be learned in any EFL syllabus, they still cause numerous difficulties for advanced EFL learners.

In relation to spatial uses of *in*, *on*, and *at*, there appears to be great confusion regarding dimensions such as:

- Leo is at the corner.* [one-dimensional: point/intersection]
Leo is standing on the sidewalk. [two-dimensional]

Leo is in the house. [three-dimensional]
Leo is at the door. [in the general area]

In addition, in all of the examples above, the Portuguese preposition *em*, would be the equivalent English preposition if the students were to construct these sentences in Portuguese. Thus, besides the dimensionality confusion presented above, the students are faced with the multiplicity of forms that the English prepositional inventory comprises compared to the Brazilian Portuguese one.

Regarding temporality, which was not as problematic as spatiality, there appears to be confusion in the following contexts:

- a) Describing parts of the day: **... the jewels paid by my father in the same night...*
- b) Longer periods: **At the 60's...*
- c) Expressions with no preposition: **... at least once in a week.*
- d) Fixed expressions: **... so, in the moment that you find...; *... because the person can enjoy their work, and in the same time...*

Research question 5: Do the advanced learners from UFSC differ significantly as to the distribution of their errors over various categories compared to the advanced learners of PUC/SP, UFU, and UNICSUL?

Table 24 displays the distribution of errors among eleven error categories across the two corpora:

Error Category	No. of errors		No. of errors		Total	%
	UFSC	%	PUC / UFU / UNICSUL	%		
LS	22	8	45	16	67	2
LP	13	5	21	7	34	1
XVPR	23	8	35	12	58	2
XNPR	10	4	26	9	36	1
XADJPR	3	1	7	3	10	
XVCO	6	2	3	1	9	
XNCO	3	1	15	5	18	
XADJCO	0	0	3	1	3	
XPRCO	1	1	10	4	11	
WR	10	4	13	4	23	
WM	3	1	11	4	14	
Total	94	34	189	66	283	100

Table 24 - Distribution of error categories across the two corpora

Two statistical tests were employed in order to find if the difference in the distribution of errors over various categories was significant or not. The first test employed was chi-square. Since the two samples differed greatly in terms of tokens and number of essays (UFSC = 13534 tokens in twenty-four essays; PUC/SP, UFU, and UNICSUL = 20220 tokens in forty-three essays) both corpora were normalized for 13.000 tokens. The p-value for the chi-square test (.2271) indicates that there is no statistically significant difference for the overall distribution of errors across the two corpora (when the values are normalized per 13.000 words).

Thus, even though these two groups of subjects may be exposed to different course syllabuses, which in turn may prioritise different grammatical features, they do not differ significantly as to the distribution of their prepositional errors over the various error categories proposed in this study.

Research question 6: Are the most frequent prepositions - *of*, *in*, *to* - the most frequently misused?

The answer to this question is twofold. If we take into account only the total number of errors per preposition, disregarding their relation to their total frequencies – the answer is *yes*. Altogether, these three prepositions accounted for 64 per cent of the total number of errors (180 errors).

On the other hand, if we construct a ratio for the incorrect uses of a preposition in relation to its total number of tokens, the answer to the research question posited above will certainly be *no*. The support for this argument can be found in Table 21, where it can be noticed that the prepositions *to*, *at*, *by*, and *on*, even not being the most frequent prepositions, were the ones with which the subjects had the greater proportions of errors.

Research question 7: Contrasted to EFL corpora of different language backgrounds, does the present corpus display similar frequencies of prepositions? Which prepositions are significantly under- and overused compared to the LOCNESS?

Table 25 provides the answer for the first question. As can be noticed, the use of prepositions across native, Brazilian, French, Spanish, Czech, Polish, Dutch, and Finnish corpora of roughly the same size shows great resemblance in terms of overall frequencies. Nevertheless, despite the great similarity in terms of overall frequencies, some prepositions display great differences in frequency across the corpora. For instance, an in-depth study would probably shed some light on why the preposition *about* occurs twice as often in the Br-ICLE as in the LOCNESS, and why the Br-ICLE displays the lowest frequencies for the prepositions *of* and *for*.

Table 25 – Frequency of prepositions across different corpora

Preposition	Locness	Br-icle	Fre	Spa	Czech	Pol	Dutch	Finn
About	45	112	56	85	116	67	69	55
At	84	84	119	94	104	119	141	72
By	211	119	134	167	113	190	216	127
For	309	268	308	300	322	319	335	324
From	118	84	138	97	101	135	151	84
In	701	717	631	808	532	694	792	788
Of	1342	935	1124	1186	990	1237	1136	1183
On	171	164	192	123	170	187	182	161
To	329	280	286	269	211	332	295	327
With	209	167	174	195	193	181	225	151
Total	3519	2930	3162	3324	2852	3461	3542	3272
Tokens	33750	33754	33241	34061	33933	33911	33132	33611

Table 26 displays the prepositions that are significantly under- and overused in the Br-ICLE compared to the LOCNESS. An alpha level of .05 ($df = 1$) was chosen for both chi-square (X^2) and Ted Dunning's Log Likelihood (G^2) tests. Any value above the critical value of 3.841 indicates that we can be 95% confident that a preposition is being significantly over

or underused. Thus, as indicated in Table 26 below, the preposition *about* is significantly overused, whereas the prepositions *by*, *from*, *of*, *to* and *with* are significantly underused.

Preposition	LOCNESS	Br-ICLE	X ²	G
About	45	112	*28,58	*27,5
At	84	84	0.00	0.0
By	211	119	*25,64	*26.0
For	309	268	2.92	2.9
From	118	84	*5.72	*5,7
In	701	717	0.20	0.1
Of	1342	935	*72.08	*73.1
On	171	164	0.15	0.1
To	329	280	*3.95	*3.9
With	209	167	*4.70	*4.7

* values are significant at $p < .05$

Table 26 – Overuse and underuse profile

It can be noticed in Table 8 that the Brazilian learners underuse most prepositions, the exceptions being *about* and *in*. It may be argued that for most learners, prepositions represent a serious trap, as their use is a matter of collocational competence and experience rather than logic. Avoiding constructions in which they are uncertain and likely to make mistakes seems only natural for them.

The ten prepositions which were under scrutiny constituted 8.38% of the tokens I have used for the investigation in the present study. (Total number of words 33754; Total number of prepositions 2930). Overall, the error rate was very low (9.65%), with a mean of 4.22 errors per essay, and an average of 43 prepositions per essay.

In the next chapter, I shall discuss the pedagogical implications and limitations of the present study, as well as indicate how it may have contributed to the fields of Corpus Linguistics, EA, and CA.

CONCLUSION

In the previous chapter, major patterns of correct and incorrect use regarding the ten prepositions under investigation were described. Some of these patterns could probably be achieved by introspection or intuition. However, personal introspection may be related to ideas about language rather than facts of it. Corpus-driven research can tell us what the facts are, and we should look for ways of deriving a theory from the data.

The collection and study of corpora of interlanguage are powerful and necessary requirements for the understanding of the production, and therefore the communication needs, of EFL learners. From such research follows the obligation to find ways of providing language learners with timely and comprehensible access to the great amount of information we are discovering about language. The way we teach today may be enhanced by conveying to the learners our growing understanding of language features, and by providing learners full access to, or significant experience with, the patterns of target language use and how particular features of their own production *deviate* from these patterns.

In the present study, the ten most frequent prepositions in the Br-ICLE were chosen in an attempt to show how, and possibly why, they *deviate* from target language patterns. A multidisciplinary approach encompassing the tools and overall rigor of Corpus Linguistics, EA, and CA, was judged adequate as the methodological framework to provide the necessary theoretical foundations for this analysis. In the remainder of this chapter I shall argue, in general terms, how my findings may have contributed to the teaching and learning of EFL, especially regarding the prepositions under investigation.

The use of corpus data in the foreign language classroom, and more specifically of concordances, may entail a change in the focus of teaching towards form, and the introduction of a more inductive approach in the learning and teaching of languages than has been the

common practice in recent years. It is undeniable that the communicative approach has helped improve learners' fluency, but it may be questioned if this strategy of focusing exclusively on meaning and overall success of communication has not overlooked the issue of accuracy.

Concordancing, or the use of computer-generated concordance lines to get students to explore regularities of patterning in the target language, may call for learners' inductive skills, or as Tribble and Jones (1990, p.20) so aptly put it, "favour learning by discovery"²⁹.

Up to now, concordancing or data-driven learning has almost solely made use of native English data. The genuineness of the data ensures that learners are presented with stretches of language which reflect the way people actually write or speak. Nonetheless, native data fails to provide corrective feedback that has shown to have positive effects on SLA. In the context of essay writing, it is learner data which can tell us what is difficult for learners in general. It would, therefore, appear to be profitable to use concordancing for learner data. The advances in computer technology have made it possible to store enormous quantities of learner data on computer and make concordances of error-prone features. In the case of the present study, the corpus was error-tagged allowing for the retrieval of different prepositional error categories in a matter of seconds. These concordances can then be presented to learners in parallel with native concordances to make learners aware of grammatical, lexical, or stylistic features, which allow for a distinction between their interlanguage and the target language.

The computerization of learner data allows for a more systematic account of learner difficulties. By inserting error tags (in this study prepositional error tags) in the learner corpus, it was possible to retrieve comprehensive lists of errors typical of a given learner population. A possible pedagogical use for these lists is to design exercises based on parallel NS/NNS concordances. For instance, in the following exercise the learners are asked to compare the prepositions that follow the noun *knowledge* in NS and NNS data.

²⁹ For a thorough discussion of induction in the learning process and concordancing in the classroom see Todd (2001), Flowerdew (1993), and Milton (1998).

Sample exercise

Compare the example from native and non-native speaker writing given below.

1. Which preposition(s) follow the noun 'knowledge'?
2. Are there prepositions which only appear in the non-native speaker examples or only in the native speaker examples? If this is the case, check whether the students are using an acceptable preposition.

(Check in your dictionary for useful information on the preposition(s) which are generally used after the noun 'knowledge')

Native speaker writing³⁰

broadening one's own knowledge	>about oneself.
reflect a knowledge	>of a different "mother tongue".
thirst for God, for the knowledge	>of why we are here
Knowledge	>of some things is beyond
his experiences and new knowledge	>of the world enable him
Argos with no knowledge	>of the town or its people,
benefiting from their knowledge	>of sexual education.
the knowledge	>of the truth of life
This new knowledge	>of genetics is leading to a
with the knowledge	>of the absurd, Caligula knows
Fear and knowledge	>of their guilt.

Non-native speaker writing³¹

the missing knowledge	> about grammar.
factors that limit the knowledge	> about the real word.
a series of esoteric knowledge	> about spiritual things,
a person has to have a great knowledge	> in his area,
professionals with good knowledge	> in foreign languages,
besides of giving you the knowledge	> of the profession,
computation and knowledge	> of English

This type of exercise may help learners become aware of the fact that the most likely preposition to follow the noun *knowledge* is the preposition *of*, with marginal cases where the preposition *about* may be suitable as well. They might ask themselves, after going through an exercise like the one presented above, if the preposition *in* is a good choice as a collocate of *knowledge*.

³⁰ LOCNESS – 182.318 tokens. The noun *knowledge* occurred 47 times and the examples display the instances where *knowledge* was followed by a preposition.

³¹ Br-ICLE – 33.754 tokens. The noun *knowledge* occurred 15 times and the examples display the instances where *knowledge* was followed by a preposition.

Exercises of this kind should be particularly motivating for learners, as they have to do with their own attested difficulties. They constitute a positive way of giving corrective feedback, since they display not only erroneous uses, but also structures that the learners have already mastered (such as the use of *about* and *of* following the noun *knowledge*). Furthermore, they combine three theoretical paradigms, namely, Corpus Linguistics, Error Analysis, and Contrastive Analysis.

If learners have access to appropriate corpus material, there is enormous scope for their own investigation of the role that prepositions play in discourse organization and the sorts of problems that learners face when making appropriate prepositional choices.

If the idea of doing corpus-based linguistic research caught on in our institution, which I personally think it could, I can envisage different types of corpora which could bring fruitful research for at least three areas of study: Translation Studies, Literature, and SLA. Here is a list of different corpora which I believe could start to be collected in the near future:

- a) A corpus of English translations into Portuguese – together with the corpus of the English original. This could be divided into several mini-corpora, depending on the nature of the texts involved (academic, literary, scientific, economic, etc).
- b) A few parallel corpora along the lines of the ICLE – any analyses carried-out on this basis would supplement and verify the conclusions derived from the study of the Portuguese ICLE.
- c) A system of comparable corpora for monitoring students' progress throughout their subsequent years of study and for 'measuring' the level of advancement at each particular year. If for each academic year, even an untagged corpus of, for example, 300-word expository essays was collected for years I, II, III, and IV separately, the amount of material on which to draw diachronic and synchronic comparisons would be abundant.

- d) A special corpus of examination papers (this would have to be carefully thought out, since most examinations take place in the classroom without computers).
- e) A separate corpus of Literature papers (especially for genre analysis – to be contrasted with other kinds of academic writing).
- f) A monitor corpus of students' essays from the British and the American varieties.
- g) Specialized subcorpora of examinations that received a common grade: A, B, C.
- h) Difficult, but also very fruitful, a corpus of attested spoken language.

The primary purpose of the error tagging procedure adopted in this study was to allow retrieval and facilitate further study of prepositional errors for advanced students, rather than to attempt a complete explication by means of the tags. It should be noted that the determination of error becomes difficult when the semantic or pragmatic intention of the writer is not clear or the syntax or lexis is so entangled that the most heroic measures cannot disambiguate meaning; an especially common problem among weak writers (e.g. **...receiving bad news like wars, misery, politic problems, and others, we are seeing how big is it and how different for culture context.*). The tags act primarily as lookups, where the most obvious and localized prepositional error types can be listed for more detailed analysis. This allows a ranking of prepositional error categories according to their absolute and relative rates of frequency, as well as according to their rate of dispersion (i.e., how frequently they occurred in separate essays).

I make no claim of having thoroughly dealt with the methodological and theoretical issues associated with error analysis techniques, such as problems in the precise identification and description of prepositional errors. In addition, up to now, I have not found a single error study which has dealt exclusively with prepositional errors which, in turn, makes the present study very exploratory and the first one of its kind to the best of my knowledge. Although these problems have been partly to blame for the decline in the academic interest in the area

of EA, they should not prevent us from exploring various linguistic learnability issues so that we have some basis for addressing discrepancies between the learners' interlanguage and the target language. Because my interest was in providing whatever feedback would be most useful to the learners, in most cases only the surface features of prepositional errors in the Br-ICLE corpus were tagged. Clearly, however, many errors have difficult-to-resolve theoretical and instructional issues that ultimately need to be addressed. Because the tagging was intended as far as possible to be theory-neutral, these features can be retrieved and explored when time permits, and references to apparently transferred Portuguese structures can be made when possible and useful. Besides corpus size, what has been stated in this paragraph constitutes, in my opinion, the major limitation of the present study.

In some ICLE studies (especially Granger, 1998), the researchers seem to have gotten carried away with their findings on the ICLE. As a result, their analysis seem to be more critical than linguistic. They conclude that the non-native students use more of what should be used less, that they do not use enough of what should be used more, that they write as if they spoke, that they ask too many questions in texts and so on – the list is endless. In other words, the learners sometimes seemed to be viewed as people of dull minds, even if in a neutral, impersonal and formal way (exactly as academic writing requires). The researchers seemed to have forgotten the fact that the learners are – learners.

It is important to remember that the learners' achievements are in accordance with the input and feedback they are given. In the case of EFL teaching, there is nothing like a compromise between quality and quantity. If we are to expect the learners' production to get closer to that of the native students, higher standards of teaching – both in terms of methodology and quality of teachers – must be combined with everyday use of English and alternatively with data-driven learning as a means to raise language awareness. We must never forget that usually the learners' results directly reflect the quality of their teaching.

Every study on ICLE must be taken not only as a compilation of statistical findings in a descriptive text, but also (and above all) as learners' feedback to their teachers and textbook writers. Through our better understanding of students' problems we can become better teachers – it is the learners who can teach us to teach them better.

REFERENCES

- Abbas, A.K. (1995). Contrastive Analysis: Is it a Living Fossil? *IRAL* Vol. 33 No. 3: 195-215.
- Abbot, G. (1980). Towards a more rigorous analysis of foreign language errors, *IRAL*, Vol. 28 No. 2:121-34.
- Bathia, A.T. (1974). An error analysis of Students Compositions. *IRAL* Vol. 12 No. 4: 337-50.
- Bell, R.T. (1974). Error Analysis: a recent pseudoprocedure in applied linguistics. *IRAL* Vol. 25-26: 35-49.
- Bley-Vroman, R. (1993). The comparative fallacy in interlanguage studies: the case of systematicity. *Language Learning* Vol. 33: 1-17.
- Brown, H. D. (1994). *Principles of Language Learning and Teaching*. Printice-Hall Regents.
- Buteau, M.F. (1970). Student's errors and the learning of French as a Second Language, A Pilot study. *IRAL* Vol. 8 No. 2: 133-45.
- Carrol, J.B. (1968). Contrastive Analysis and Interference Theory, in *Monograph Series on languages and linguistics* Vol. 21. James E. Atlantis (ed.). In Bathia, A.T. (1974). An error analysis of Students Compositions. *IRAL* Vol. 12 No. 4: 337-50.
- Celce-Murcia, M. & Larsen-Freeman, D. (1983). *The Grammar Book: An ESL/EFL Teacher's Course*. Newbury House Publishers, Inc.
- Celce-Murcia, M. & Larsen-Freeman, D. (1999). *The Grammar Book: An ESL/EFL Teacher's Course*. Heinle & Heinle Publishers.
- Colling Cobuild English Guides 1: Prepositions*
- Corder, S.P. (1967). The significance of learner's errors. *IRAL*, Vol. 5 No. 4: 161-70.
- Corder, S.P. (1971). Idiosyncratic dialects and error analysis. *IRAL* Vol. 9 No. 2: 147-60.
- Corder, S.P. (1975). Error analysis, Interlanguage and second language acquisition (Survey Article)., *Language teaching and Linguistic Abstracts* Vol. 8: 201-18, in James, C. (1998). *Errors in Language Learning and Use: Exploring error Analysis*. Longman, Longman & New York.
- Dagneaux, E., Denness, S., Granger, S. (1998). Computer-aided Error Analysis, *System* Vol.26 No. 3:163-74.
- Dagneaux, E., Denness, S., Granger, S., & Meunier, F. (1996) *Error Tagging Manual Version 1.1*. Centre for English Corpus linguistics, Université Catholique de Louvain, Louvain-la-Neuve.
- Dunning, T. (1993). Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics*, Vol.19 No.1: 61-74.
- Duskova, L. (1969). On sources of errors in foreign language learning. *IRAL*, Vol.7 No.1: 11-36.
- Flowerdew, J. (1993). Concordancing as a tool in course design. *System*, Vol.21:231-44.

- Flowerdew, L. (1998). Integrating 'Expert' and 'Interlanguage' Computer Corpora Findings on Casuality: Discoverics for Teachers and Students. *English for Specific Purposes*, Vol.17:329-45.
- Grabowski, J. & Weiss, P. (1996). The prepositional inventory of languages: a factor that affects comprehension of spatial prepositions. *Language Sciences*, Vol.18:19-35.
- Granger, S. (ed.) (1998). *Learner English on Computer*, London & New York.
- Granger, S. (1992). *International Corpus of Learner English*. In: Papers from the Thirteenth International Conference on Language Research on Computerized Corpora, Nijmegen 1992. Edited by Jan Aarts, Pieter de Haan and Nellake Ostdijk. Rodopi B.V., Amsterdam-Atlanta, GA. 57-72.
- Granger, S. (1998). The computer learner corpus: a versatile new source of data for SLA research. In: Granger, S. (ed.) (1998). *Learner English on Computer*, London & New York.
- Granger, S. & Tyson, S. (1996). Connector usage in English essay writing of native and non-native speakers of English. *Word Englishes*, Vol.1:17-27.
- Greenbaum, S. (1988). *Good English and the Grammarian*. Longman, London.
- Grimm, H. (1975). On the child's acquisition of semantic structure underlying the wordfield of prepositions. *Language and Speech*, Vol.2:97-119. In Tomasello, M. (1987). Learning to use prepositions: a case study. *Journal of Child Language* Vol. 14: 79-98.
- Hammarberg, B. (1974) The insufficiency of error analysis. *IRAL*, Vol. 12: 185-92.
- Hawkins, J.A. (1987). Implicational universals as predictors of language acquisition. *Language Learning* Vol. 25: 453-73.
- Hughes, G.A. & Lascaratou, C. (1982). Competing criteria for error gravity. *English Language Teaching Journal* Vol. 36 No. 3: 175-82.
- James, C. (1994) Don't shoot my dodo: on the reliance of Contrastive and Error Analysis. *IRAL*, Vol. 32 No. 3: 179-200.
- James, C. (1998). *Errors in Language Learning and Use: Exploring error Analysis*. Longman, Longman & New York.
- Johnston, J. & Slobin, D. (1979) The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language* Vol. 6: 529-46.
- Kasper, G. & Kellerman, E. (1997). Introduction, in G. Kasper and E. Kellerman (eds), *Communication strategies: Psycholinguistic and Sociolinguistic Perspectives*. Addison Wesley Longman, London. pp 1-13.
- Kilgarriff, A. (1997). *Using word frequency lists to measure corpus homogeneity and similarity between corpora*. Proceedings 5th ACL workshop on very large corpora. Beijing and Hong Kong.
- Kilgarriff, A. (2001). *Comparing Corpora*. To appear in the International Journal of Corpus Linguistics (2001).
- Lennon, P. (1991a). Error and the very advanced learner. *IRAL* Vol. 39 No. 1: 31-44.
- Lennon, P. (1991b). Error: some problems of definition, identification and distinction. *Applied Linguistics* Vol. 12 No. 2: 180-96.

- Lindstromberg, S. (2001). Preposition entries in UK monolingual learners' dictionaries: Problems and possible solutions. *Applied Linguistics* Vol. 22 No. 1:79-103.
- Milton, J. (1998). *Exploiting L1 and interlanguage corpora in the design of an electronic language learning and production environment*. In. Granger, S. (ed.) (1998). *Learner English on Computer*, London & New York.
- Nemser, W. (1971) Approximative systems of foreign language learners. *IRAL* Vol. 9 No. 2: 115-23.
- Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik (1985). *A Comprehensive grammar of the English Language*. London: Longman.
- Rastall, P. (1994). The preposition flux. *IRAL* Vol. 33: 229-31.
- Rayson, P. & Garside, R. (2000). *Comparing Corpora using Frequency Profiling*. In Proceedings of the workshop on *Comparing Corpora*, held in conjunction with the 38th annual meeting of the Association for Computational Linguistics. 1-8 October 2000, Hong Kong. 1-6.
- Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: a cross-linguistic approach. In. Granger, S. (ed.) (1998). *Learner English on Computer*, London & New York.
- Sardinha, T.B. (1999). Computador, discurso e lista de palavras chave: Um quadro teórico. <http://www.cursos.f2s.com>.
- Schachter, J. & Celce-murcia, M. (1977). Some reservations concerning error analysis. *TESOL Quarterly* Vol. 11 No. 4: 441-51.
- Schachter, J. (1974). An error in error analysis. *Language Learning* Vol. 24 No. 2: 205-14.
- Sciarone, A.G. (1970). Contrastive Analysis - Possibilities and Limitations. *IRAL* Vol. 8 No. 2: 115-31.
- Scott. M.R. (1996). *Wordsmith Tools*. Oxford University Press.
- Selinker, L. (1972). Interlanguage. *IRAL* Vol. 10 No. 3: 200-31.
- Stubbs, M. (1996). *Text and Corpus Analysis: Computer-assisted Studies of language and Culture*. Blackwell Publishers.
- Swan, M. (1981). *Practical English Usage*. Oxford University Press.
- Swan, M. (1997). *Practical English Usage - New Edition*. Oxford University Press.
- Todaka, Y. (1996). Between and among: a data base analysis. *Word* Vol. 47: 13-40.
- Todd, R. W. (2001). Induction from self-selected concordances and self-correction. *System*, Vol.29:91-102.
- Tomasello, M. (1987). Learning to use prepositions: a case study. *Journal of Child Language* Vol. 14: 79-98.
- Towell, R. & Hawkins, R. (1994). *Approaches to Second Language Acquisition*. Multilingual Matters, Clevedon.
- Tribble, C. & Jones, G. (1990). *Concordances in the Classroom*. Longman, London.
- Vandeloise, C. (1991). *Spatial Prepositions: a case study from French*. The University of Chicago Press, Chicago and London.

- Voster, J. (1984). The first prepositions in Afrikaans: order and semantic distinctions. Paper presented to the Third International Congress for the study of Child language, Austin, Texas. In Tomasello, M. (1987). Learning to use prepositions: a case study. *Journal of Child Language* Vol. 14: 79-98.
- Wardhaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly* Vol. 4 No. 2: 123-30.
- Watkins, R. V. & Rice, M. L. (1991). Verb Particle and Preposition Acquisition in Language-Impaired Preschoolers. *Journal of Speech and Hearing Research* Vol. 34: 1130-1141.

APPENDIX A

Word List - Frequency of first 100 words

N	Word	Freq.	%	N	Word	Freq.	%
1	THE	1.788	5,30	51	LOVE	97	0,29
2	TO	1.192	3,53	52	MANY	93	0,28
3	AND	949	2,81	53	HAS	88	0,26
4	OF	935	2,77	54	WHEN	87	0,26
5	A	853	2,53	55	OTHER	85	0,25
6	IS	776	2,30	56	AT	84	0,25
7	IN	717	2,12	57	FROM	84	0,25
8	THAT	563	1,67	58	SOCIETY	83	0,25
9	IT	434	1,29	59	THINGS	82	0,24
10	ARE	392	1,16	60	WAY	79	0,23
11	THEY	336	1,00	61	GET	76	0,23
12	NOT	328	0,97	62	HIS	76	0,23
13	HAVE	311	0,92	63	PERSON	76	0,23
14	PEOPLE	301	0,89	64	TIME	76	0,23
15	BE	279	0,83	65	WORLD	76	0,23
16	FOR	268	0,79	66	WHICH	75	0,22
17	THIS	256	0,76	67	HER	74	0,22
18	MORE	215	0,64	68	JUST	73	0,22
19	AS	212	0,63	69	THESE	73	0,22
20	YOU	202	0,60	70	VERY	73	0,22
21	WE	200	0,59	71	SHE	71	0,21
22	THEIR	186	0,55	72	SHOULD	71	0,21
23	OR	178	0,53	73	ONLY	70	0,21
24	BUT	177	0,52	74	THAN	70	0,21
25	CAN	177	0,52	75	WOULD	70	0,21
26	I	168	0,50	76	ALSO	69	0,20
27	WITH	167	0,49	77	THINK	69	0,20
28	ON	164	0,49	78	IMPORTANT	68	0,20
29	IF	154	0,46	79	LIKE	68	0,20
30	WHAT	154	0,46	80	IT'S	66	0,20
31	DO	137	0,41	81	MOST	65	0,19
32	ALL	132	0,39	82	EVEN	63	0,19
33	BECAUSE	127	0,38	83	LIVE	62	0,18
34	THERE	125	0,37	84	TELEVISION	60	0,18
35	SO	124	0,37	85	BEING	59	0,17
36	ONE	122	0,36	86	CHILDREN	59	0,17
37	SOME	122	0,36	87	HOW	59	0,17
38	BY	119	0,35	88	MAKE	58	0,17
39	MONEY	118	0,35	89	NO	58	0,17
40	ABOUT	112	0,33	90	SOMETHING	58	0,17
41	OUR	112	0,33	91	COULD	56	0,17
42	AN	110	0,33	92	NOWADAYS	56	0,17
43	WAS	109	0,32	93	UNIVERSITY	56	0,17
44	LIFE	108	0,32	94	WERE	55	0,16
45	THEM	106	0,31	95	TV	54	0,16
46	WHO	106	0,31	96	ORDER	53	0,16
47	HE	105	0,31	97	MUCH	52	0,15
48	WILL	105	0,31	98	YOUR	51	0,15
49	JOB	99	0,29	99	MY	50	0,15
50	GOOD	97	0,29	100	REALLY	50	0,15

APPENDIX B

Print Screen shot of WordSmith Wordlist - Statistics

The screenshot shows the WordSmith Wordlist Statistics window. The window title is 'WordSmith [new Wordlist (S)]'. The menu bar includes 'File', 'Settings', 'Comparison', 'Index', 'Window', and 'Help'. The toolbar contains various icons for file operations and search. The main area displays a list of words and their frequencies, sorted in descending order. The words are listed in a column on the left, and their corresponding frequencies are listed in a column on the right. The list includes words like 'UFSC-PUC.TXT', '193,776', '33,754', '4,142', '12,27', '40,64', '4,56', '1,104', '25,43', '19,16', '166', '203,34', '246,82', '0', '1,065', '6,779', '6,624', '6,259', '3,578', '2,890', '2,151', '1,601', '1,348', '825', '406', and '233'.

Word	Frequency
UFSC-PUC.TXT	193,776
	33,754
	4,142
	12,27
	40,64
	4,56
	1,104
	25,43
	19,16
	166
	203,34
	246,82
	0
	1,065
	6,779
	6,624
	6,259
	3,578
	2,890
	2,151
	1,601
	1,348
	825
	406
	233

APPENDIX C

Sample of concordance lines of *at*

	File	Line	Word	Count	File	Line
		dre Herchcovitch who was graduated in Fashion at the Faculdade Santa Marcelina. Now Santa	at	124	31.txt	78
		e are living in fear, mainly in the big cities. Here, at Fpolis, violence is not so strong, but it has in	at	59	ne.txt	11
		iage is a farce only at the time of your grandma. At that time women were obligated to marry unt	at	33	25.txt	12
		e. A good and harmonious relationship at home, at school, at work and in everywhere happens	at	441	arli.txt	80
		unk-drivers (and probably most of us had heard, at least, something similar). Last year, one of t	at	185	10.txt	43
		Since the beginning of the humanity history, at least the history that is studied at the school	at	13	icio.txt	2
		assurance: job's future, as we know it today, is at a check-mate position, aggravated by the wo	at	115	39.txt	29
		cts in spray maybe they will not be produced, if at least we change habits, maybe things will ch	at	255	101.txt	49
		w to transfer facts and situations into imagines. At last but not least, it is important to say that i	at	343	ne.txt	71
		er think about her life, what was more important at that time, how she would like to be, what wer	at	187	21.txt	66
		y, different schedule and etc. But when family is at home, every one often spend the free time in	at	155	ara.txt	59
		. In the middle east they have been fighting for at least 2,000 years. Though it's said that it's a	at	394	icio.txt	51
		lems is better than having job insecurity, not job at at all.	at	663	15.txt	100
		personal accomplishment to children, it looks - at least to me - very much like a parent. If childr	at	457	101.txt	73
		etween 1992 and 2000 will demand knowledges at maths, reading comprehension, as well as a	at	253	39.txt	64
		nals with good knowledge in foreign languages, at least two, university degree, and also, habilit	at	130	37.txt	46
		sesperate because nowadays their relative look at them as a trouble and send them to asylum.	at	290	27.txt	58
		great-great-grandchildren. They are also looking at our genes, which they believe hold the key to	at	393	15.txt	59
		iety. One can easily recognize that, just looking at benefactions which are made and used by a	at	408	102.txt	67
		very strong feeling - some times it is called "love at first sight". The feeling is the first to build a	at	107	24.txt	32
		ur. On the other hand, all of them could mention at least nine serial killers. The world has chang	at	567	ina.txt	91
		e financial security than staying without any job at all. This case worth also for the case of peo	at	471	36.txt	51
		r the water, and stay there for two/three minutes at least; and, in "My Fair Lady", to be consider	at	303	38.txt	31
		friends. Frequently, a child has only one parent at home, who is often absent, working. The chil	at	181	101.txt	31
		o many beggar, so many kids asking for money at the lights, so many thieves. Not because the	at	356	ulia.txt	68
		e stores and buy things that they aren't needing at the moment, just to be in faction. Even thou	at	404	bio.txt	66
		ific objective. It is important to set one objective at a time. After reaching the first one, then it is	at	130	38.txt	14
		ed before. Say that the marriage is a farce only at the time of your grandma. At that time wome	at	27	25.txt	10
		re-son-the data				

APPENDIX D

Sample of concordance lines of *to* tagged as a preposition using the TOSCA tagger

Line	Concordance	Count	Tag
3571	ding_VB(lex.montr.ingp) the_ART(def) mass_N(sing) to_PREP(ge) thing_N(sing) according_PREP(ge):1/2	11.107	oi
3572	u) are_VB(aux.pass.pres) related_VB(lex.montr.edp) to_PREP(ge) the_ART(def) acquisition_N(sing) of_PR	13.527	oi
3573) from_PREP(ge) B_NUM(card.sing) % ADV(ge.pos) to_PREP(ge) 90_NUM(card.sing) % ADV(ge.pos) on	16.278	oi
3574	pres) the_ART(def) main_ADJ(ge.pos) target_N(sing) to_PREP(ge) companies_N(plu) advertisements_N(pl	19.648	oi
3575	B(aux.modal.pres) be_VB(lex.intr.infin) living_N(sing) to_PREP(ge) the_ART(def) age_N(sing) of_PREP(ge)	11.399	oi
3576	res) the_ART(def) original_ADJ(ge.pos) words_N(plu) to_PREP(ge) " _PUNC(oquo) All_PRON(univ) men_N(37.447	oi
3577) be_VB(aux.pass.infin) ascribed_VB(lex.montr.edp) to_PREP(ge) the_ART(def) American_ADJ(ge.pos) lo	6.801	oi
3578	past) be_VB(lex.cop.infin) a_ART(indef) limit_N(sing) to_PREP(ge) the_ART(def) amount_N(sing) of_PREP	3.258	oi
3579	board) you_PRON(pers.number) talk_VB(lex.intr.pres) to_PREP(ge) boys_N(plu) and_CONJUNC(coord) girl	4.329	oi
3580	_PUNC(comma) to_PRTCL(to) travel_VB(lex.intr.infin) to_PREP(ge) France_N(sing) and_CONJUNC(coord)	28.086	oi
3581	(lex.cop.pres) financial_ADJ(ge.pos) burden_N(sing) to_PREP(ge) society_N(sing) and_CONJUNC(coord)	31.878	oi
3582	is_VB(aux.pass.pres) destined_VB(lex.montr.edp) to_PREP(ge) something_PRON(ass) and_CONJUNC	21.961	oi
3583	RON(rel) is_VB(lex.cop.pres) important_ADJ(ge.pos) to_PREP(ge) you_PRON(pers.number) as_PREP(ge)	6.973	oi
3584	VB(lex.cop.past) never_ADV(neg) able_ADJ(ge.pos) to_PREP(ge) became_VB(lex.cop.past) a_ART(indef)	30.077	oi
3585	not_ADV(neg) give_VB(lex.montr.infin) names_N(plu) to_PREP(ge) them_ADJ(ge.pos) A_ART(indef) pers	26.659	oi
3586	poor) for_PREP(ge) poor_ADJ(ge.pos) families_N(plu) to_PREP(ge) paying_VB(lex.montr.ingp) a_ART(indef	36.709	oi
3587	agers_N(plu) going_VB(lex.intr.ingp) out_ADV(phras) to_PREP(ge) a_ART(indef) bar_N(sing) _PUNC(com	30.565	oi
3588	ast) have_VB(aux.perf.infin) liked_VB(lex.montr.edp) to_PREP(ge) _PUNC(per) Based_VB(lex.montr.pa	13.116	oi
3589	os) taking_VB(lex.montr.ingp) them_PRON(pers.plu) to_PREP(ge) the_ART(def) beauty_N(sing) parlor_N(7.140	oi
3590	(sup) one_PRON(one) compared_VB(lex.montr.edp) to_PREP(ge) them_ADJ(ge.pos) Because_CONJU	23.365	oi
3591	ere_VB(lex.cop.past) just_ADV(ge.pos) tools_N(plu) to_PREP(ge) nobles_N(plu) become_VB(lex.cop.infin	29.161	oi
3592	make_VB(lex.montr.infin) the_ART(def) path_N(sing) to_PREP(ge) employment_N(sing) become_VB(lex.c	26.067	oi
3593	(ge.pos) _PUNC(per) Backing_VB(lex.montr.ingp) to_PREP(ge) the_ART(def) begining_N(sing) of_PRE	28.233	oi
3594	,sing) communism_N(sing) applies_VB(lex.intr.pres) to_PREP(ge) human_ADJ(ge.pos) beings_N(plu) _P	3.240	oi
3595	,pres) been_VB(lex.intr.edp) dangerous_ADJ(ge.pos) to_PREP(ge) the_ART(def) body_N(sing) and_CONJ	36.849	oi

APPENDIX E

Sample of concordance lines bearing the error code (XVPR) using *WordSmith Concord*

Concordance

1	things were easier and other were more difficult to be learned, it (XVPR)	depends of \$depends on\$ the abilities of each one. W
2	out each other with real love and without personal interest. It (XVPR)	depends of \$depends on\$ human spiritual evolution. F
3	involving money, just as the embezzlement of money which is (XVPR)	destinated to \$destinated for\$ something and some poli
4	P) the rhythm of each other \$each other's rythm\$, they have to (XVPR)	devote themselves on \$devote themselves to\$ this plot.
5	he has to wear the right clothes to fit in society. Clothes stores (XVPR)	direction their products to \$direct their products for / to
6	e wants to feel save being sure that his mother was there. Rita (XVPR)	enrols at \$enrols in\$ a program called Open University.
7	NPR) perspective of \$perspective on\$ life inside society and not (XVPR)	excluded of \$excluded from\$ it. As we could see a so
8	, he had finished what he could do. Now he was satisfied and (XVPR)	filled of \$filled with\$ victory. He did not want the big bo
9	ve to follow. It is not so difficult to find friends or colleagues that (XVPR)	graduated at \$graduated from\$ college and didn't guet
10	ublic university and now it is time for their graduation. What will (XVPR)	happen with \$happen to\$ them? They will have to com
11	erstand not only what happened in Russian revolution, but still (XVPR)	happening with \$happening to\$ the ideas of our societ
12	ed and needs stay alone for some minutes. It's normal and can (XVPR)	happens with \$happen to\$ everybody. The respect, co
13	ortant? Nowadays, to attend an university is really (XADJCO) (XVPR)	important for preparing yourself to \$important to prepar
14	ar by \$swear to\$ God, and besides acquitting a black rich man (XVPR)	indicted by \$indicted for\$ double homicide, they also a
15	the free time in front of the television. As a result, they do not (XVPR)	interact to \$interact with\$ each other. Rather than read
16	itation laws and human rights The government has no right to (XVPR)	interfere on \$interfere in\$ people's personal affairs. Ther
17	rocess, the way they reached their objective and what they had (XVPR)	learned with \$learned from\$ it, as their limits, fears and
18	at the schools along the country, we can clearly notice that we (XVPR)	live on \$live in\$ a constant fight for power. Much of wha
19	dual citizen does in front of this set, s/he uses television just to (XVPR)	look for \$look at\$. Obviously, this is the principal obje
20	respect and (XNCO) desire of learning \$desire to learn\$. If we (XVPR)	look to the art or to the love \$look at art or love\$ as so
21	any case exposed before, please turn off the television set and (XVPR)	look to \$look at\$ your family or take one book and read
22	he same reasons given by Eric Fromm. I have a diferente way of (XVPR)	looking into \$looking at\$ it. I really think love is an art,
23	of using\$ glasses made of plastic we should use just the ones (XVPR)	made by \$made of\$ glass or metal, if we decide not to
24	all, I believe it's fundamental to try to know the person you'll get (XVPR)	married with \$married to\$ as well as possible. There's
25	authority and who are the victims. Many of these programs are (XVPR)	oriented for \$oriented to\$ people with lower education a

re-son the data

Inicio WordSmith Tools Microsoft Word 00:00

APPENDIX F

Print screen shot of attributing an XVPR tag and the corrected form in the text using the

UCLEE

